ADAPTASIA: AN ADAPTIVE MUSIC PIECE

by

SARAH KITTEN

A THESIS

Presented to the Department of Music
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor or Science

Fall 2022

An Abstract for the Thesis of

Sarah Kitten for the degree of Bachelor of Science
in the Department of Music to be taken Fall 2022

Title: ADAPTASIA: An Adaptive Music Piece

Approved: *Brian McWhorter. Ph.D.*
Primary Thesis Adviser

ADAPTASIA is a piece of indeterminate digital music which adapts in real time to user inputs. I used an audio engine called FMOD Studio to arrange my original music into a structure which randomizes some elements of the piece and links others to user-configurable parameters. In this paper I explain the inner workings of ADAPTASIA and the user interface that I created. This thesis also includes a discussion of how ADAPTASIA relates to acoustic improvisation and indeterminate music, a brief overview of the current state of adaptive music in video games, and my reflections on the project.

Table of Contents

Link to online interface and video demos:
https://blogs.uoregon.edu/adaptasia/

List of Figures

Introduction

This paper centers on my adaptive music project, ADAPTASIA. I used an audio engine called FMOD Studio to organize short clips of original music into an algorithmic music piece with features that can be controlled by real-time user input. The structure of the piece is highly indeterminate so that it can loop indefinitely while presenting new sounds, textures, and musical ideas.

ADAPTASIA was inspired by my curiosity about the increasingly complex, modular possibilities for making video game music that dynamically reacts to the player's actions. Unlike a movie score, a video game soundtrack can't be pre-composed to align with every action on the screen because the composer cannot predict the player's decisions. The simplest way to score a video game is to write a set of looping songs that fade in and out depending on what area or level the player enters, with sound effects that play in response to important events. However, as the field of video game scoring progresses, more game soundtracks are designed as intricate systems which change their organization or instrumentation in real time based on the gameplay scenario, conveying important information to the player while achieving the smooth and immersive artistic effect of a continuous, pre-composed score.

After a technical explanation of ADAPTASIA, I will relate my project to indeterminate music designed for live performance, such as acoustic improvisation and the famously aleatoric music of John Cage. Then, I will consider history, techniques, and case studies of adaptive music for emotional immersion in video games. Finally, I'll reflect on the process, outcomes, and value of the project.

ADAPTASIA: The Project in Detail

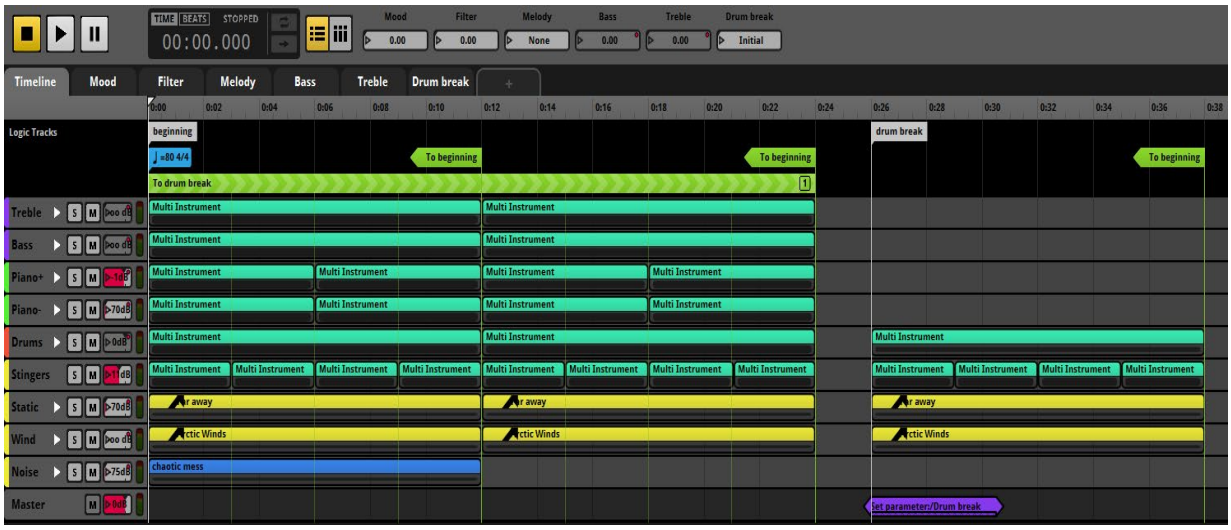

Figure 1: The full FMOD timeline of ADAPTASIA

Pictured is a view of the full FMOD timeline of ADAPTASIA. The top of the screen shows the playback controls and the values of the four user-configurable parameters that I defined for the piece - Mood, Filter, Melody, and Drum break - and the two 'helper' parameters, Bass and Treble. Beneath are tab headings, a time bar displaying seconds, and the set of logic markers for the piece, including the section label 'beginning' and two 'To beginning' transitions. The rest of the image consists of the audio tracks arranged along the timeline and labeled on the left. The labels are color-coded according to category:

- Treble and Bass: The two options for melody, determined by the 'Melody' parameter
- Piano+ and Piano-: The two options for piano accompaniment, determined by the 'Mood' parameter
- Drums: Drum accompaniment
- Stingers, Static, Wind, and Noise: Sound effects determined by the 'Mood' parameter
- Master: The master track which controls the overall gain (volume)

Structure

ADAPTASIA continues to play infinitely once it has started, but its highly modular and randomized structure keeps it from ever directly repeating on a large scale. The larger structure chooses between three four-bar phrases: one harmonized to a tonic (I) chord, one harmonized to a dominant (V) chord, and one containing a drum solo. The choice of phrase is controlled and randomized through the logic markers and transitions. The tonic phrase ranges from the 'beginning' marker to the first 'To beginning' transition. This transition is set to have a 30% chance to transition playback back to the 'beginning' marker, repeating the tonic section. I chose this probability because it led to the most satisfying-feeling proportion of tonic and dominant.

Figure 2: Settings for the first 'To beginning' transition, with Chance=30%

If the transition does not activate, the piece continues through the dominant section to the second 'To beginning' transition, which always transitions back to the 'beginning' marker. The effect is that the tonic section sometimes repeats itself, while the more unstable dominant section always relaxes back to tonic. This resembles AABA and ABA structures common in music, but introduces enough randomness to keep the simple harmonic structure from ever becoming completely predictable.

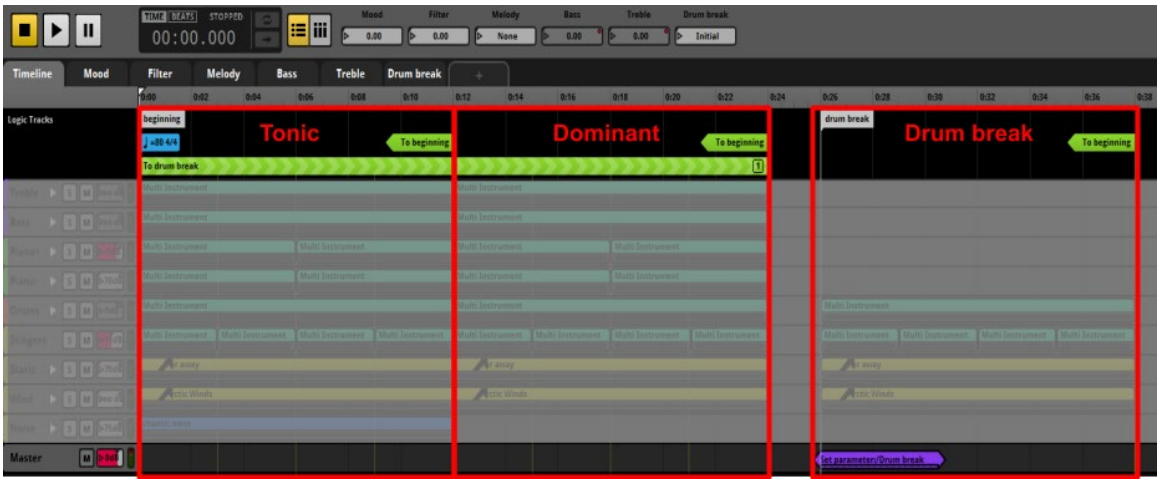

Figure 3: The section structure of ADAPTASIA: Tonic, Dominant, and Drum break

The drum break section is accessed through the user-controlled ‘Drum break’ parameter. When this parameter is set to ‘Trigger’, the ‘To drum break’ transition region that spans both the tonic and dominant sections is activated. This transition region has its quantization interval set to 1 bar, so playback will continue until the end of the current bar, then jump to the ‘drum break’ marker. A command instrument (in purple on the Master track) sets the parameter back to ‘Initial’, and when the ‘To beginning’ transition at the end of the drum break section is reached, playback returns to the beginning. In this way, the user can enact some control over the structure of the piece by interrupting its normal flow. See a video demo of the ‘To drum break’ parameter here: https://blogs.uoregon.edu/adaptasia/drum-break/.

Figure 4: Settings for the 'To drum break' transition region, quantized to 1-bar intervals

The interior structure of each music track provides even more randomness through the use of asynchronous multi-instruments. Each multi-instrument contains multiple audio files and plays on its own timeline after being triggered by playback. Let's examine the first multi-instrument in the 'Piano+' track - clicking on it shows the following window:

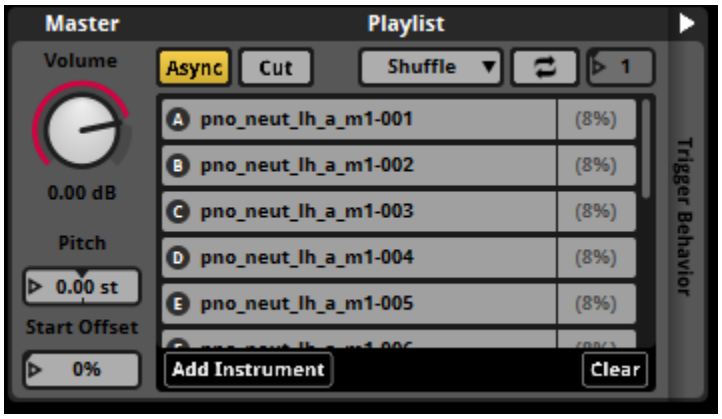

Figure 5: Settings for the first multi-instrument in the 'Piano+' track, with 13 possible audio files to play

This instrument contains thirteen audio files - when the instrument is triggered, one file is selected and played, so each has an ~8% chance of being played (100% / 13 = 7.69%). In this case, each file is a slightly different version of the first two bars of a piano phrase. The next

multi-instrument in the track contains different options for ending the phrase, and the two multi-instruments after that contain options for beginning and ending a phrase in the dominant harmony. The slight variations in the performance of the phrases serves to keep the music sounding fresh. See a video demo of the 'Piano+' multi-instruments here: https://blogs.uoregon.edu/adaptasia/piano/.

'Treble', 'Bass', and 'Drums' are grouped into longer four-bar phrases with more variety and internal structure, while the 'Stingers' consist of unsettling one-bar audio clips. 'Treble' uses a high-pitched sample-based synthesizer, 'Bass' uses a virtual instrument sampling a Rickenbacker bass, and 'Drums' uses samples from a drum kit. 'Noise' only has one audio file (a percussive, mechanical synthesized sound), so it is represented as a single instrument rather than a multi-instrument.

'Static' and 'Wind' use event instruments, which reference externally defined musical events. The event instruments for 'Static' and 'Wind' are designed similarly to each other, consisting of multiple multi-instruments with different delay intervals and the same set of synthesizer audio files, each playing a single note. A delay interval determines the range of how long an activated instrument will wait before playing an audio file - an instrument with a 5-10 second delay interval will wait between 5 and 10 seconds to select and play an audio file. Each multi-instrument in 'Static' and 'Wind' also has a randomized volume within a specified range.

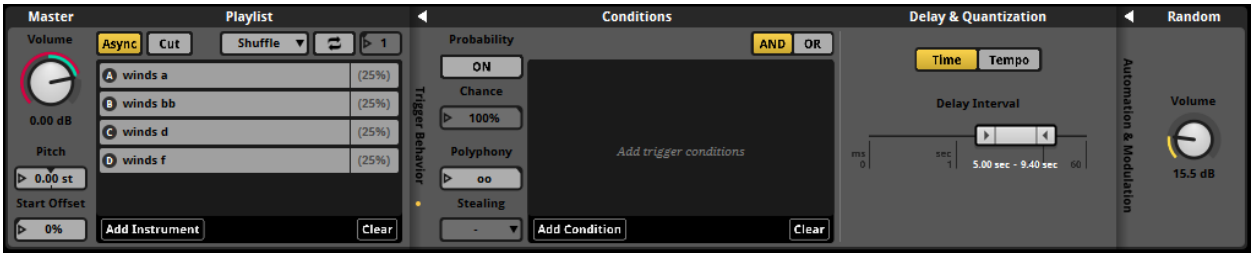

Figure 6: Settings for the 'Wind' event instrument, with a randomized note, delay, and volume

The result is a musical phrase where the notes, rhythms, and volumes are randomized within specified ranges. However, because the synthesizers for both 'Static' and 'Wind' are more ethereal than melodic, the phrases blur into an ambient overlay rather than functioning as a melody. See video demos of 'Static' and 'Wind' here: https://blogs.uoregon.edu/adaptasia/wind-static/.

Parameters

I designed four user-configurable parameters in ADAPTASIA - Mood, Filter, Melody, and Drum break - and the two 'helper' parameters, Bass and Treble. Events, probabilities, and volumes can be triggered based on the values of these parameters, transforming the content of the piece in dramatic or subtle ways. Each of the parameters with a numerical value has a set 'seek speed', which prevents it from jumping suddenly between values but instead slides smoothly between them at a given speed.

The heart of the user interaction in ADAPTASIA lies in the 'Mood' parameter, which transforms the piece from cheerful and jazzy at its lowest value to distorted and anxious as the value increases, and finally to a slow, distressed ambient track at its highest value. Most of these transformations are achieved through volume automation - certain tracks, such as the 'Treble', 'Bass', and 'Piano+' lines, fade out as the 'Mood' value increases, to be replaced by 'Piano-' and the 'Static', 'Wind', and 'Noise' sound effects. Although all the tracks are playing at the same time, the volume automation determines which can be heard based on the 'Mood' value, resulting in very different versions of the piece. The multi-instruments in the 'Stingers' track also depend on the 'Mood' parameter: each one has a 35% chance of playing when triggered, but

only if the ‘Mood’ value is at least 0.15. See a video demo of the ‘Mood’ parameter here: https://blogs.uoregon.edu/adaptasia/mood/.

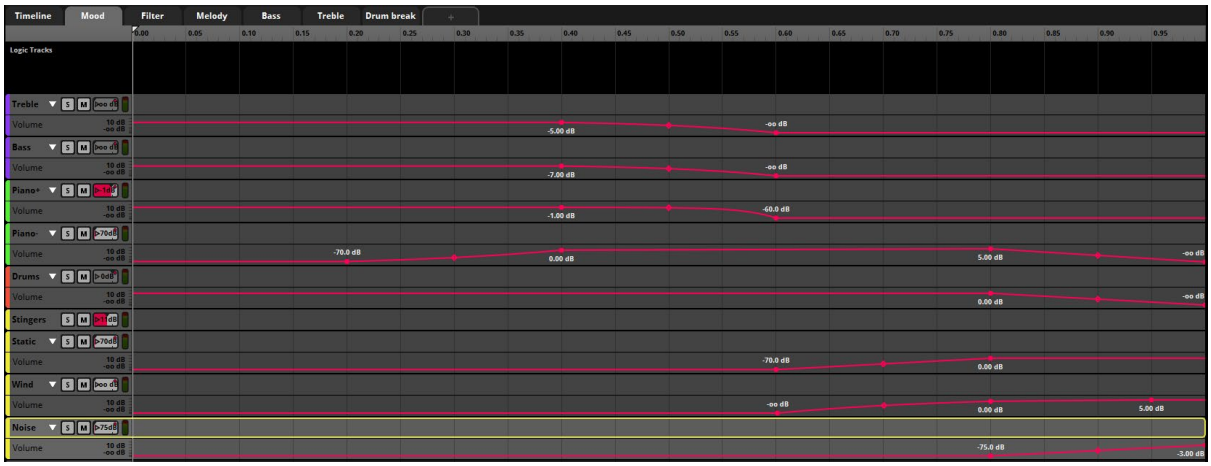

Figure 7: Volume automation for the ‘Mood’ parameter

The rest of the parameters in ADAPTASIA are considerably simpler. ‘Filter’ uses an FMOD tool called a Snapshot to fade in a set of mixer settings, which in this case includes a low-pass filter and a Chorus effect. The outcome is a dulled, slightly distorted sound, as if the listener is underwater. I chose this effect because it creates an interesting ‘emotional distance’ at all points along the ‘Mood’ spectrum, and because it is a common effect in games for underwater scenes, menus, and emotional pullbacks. See a video demo of the ‘Filter’ parameter here: https://blogs.uoregon.edu/adaptasia/filter/.

The ‘Melody’ parameter triggers which melody instrument should be heard: a synthesizer (called ‘Treble’), a bass guitar (called ‘Bass’), or neither. In order to allow for any transition among these three options, I created the helper parameters ‘Bass’ and ‘Treble’ to automate the volumes of their respective tracks. A change in melody instrument can be a simple and effective leitmotif for a character, location, or gameplay scenario in a video game, and serves as a simpler

example of effective adaptive music to display in ADAPTASIA. See a video of the 'Melody' parameter here: https://blogs.uoregon.edu/adaptasia/melody/.

Finally, as discussed in the 'Structure' section (p. 8), the 'Drum break' parameter triggers a transition into a drum solo which leads back into the main piece.

Interface

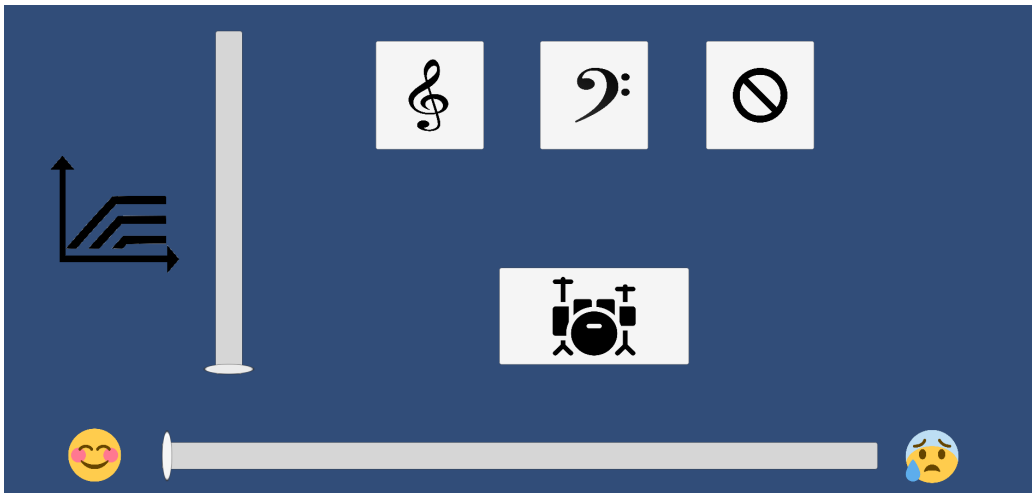

Figure 8: Web-hosted visual interface for ADAPTASIA

In order to make ADAPTASIA easily accessible and interactable, I used a video game engine called Unity to build a prototype user interface, which I uploaded online: https://blogs.uoregon.edu/adaptasia/ ("Unity" 2020). The interface is graphically simple and provides no written description of its functions - the user is encouraged to experiment with the buttons and sliders and observe their effect on the music. Each item in the interface corresponds with one of the user-configurable parameters previously described: the 'treble clef', 'bass clef', and 'none' buttons toggle among the 'Melody' parameter options, the 'drum' button triggers the quantized drum break, the 'filter' slider on the right controls the 'Filter' parameter, and the 'happy/anxious' slider on the bottom controls the 'Mood' parameter.

The interface hides the inner workings of ADAPTASIA and serves as a prototype of the kind of simple sandbox that I envision for improvisatory uses of adaptive music. Made more abstract, the interface could become a video game where the player's actions impact the parameters without direct manipulation, causing the game music to adapt seamlessly to align with gameplay.

ADAPTASIA and Improvisation

Is ADAPTASIA a composition or an improvisation? Let us examine Paul Turowski's definitions of the terms:

> Musical composition and free improvisation might be considered opposing terminals on a continuum, upon which all musical performances reside. I broadly define a composition as any pre-arranged plan of action for a performer, whether it specifies a sonic result or performance parameters such as physical position, mental state, etc. (A composition may also exist as a "fixed" medium - e.g. a tape piece - but such a composition is still performed, even if somewhat automatically.) By contrast, the act of improvisation more involves real-time musical decisions by a performer. (Turowski 2016)

ADAPTASIA exhibits traits from both composition and improvisation: it exists as a "pre-arranged" plan of action for the "performing" computer, but one that specifies results across several dimensions of user-controlled parameters. However, the user, whose influence over the music is heavily abstracted, may not necessarily be considered a "performer." If we imagine a performance where musicians spontaneously create music to match the movements of an improvising dancer, the user of ADAPTASIA fills a similar role to the dancer: influencing the sound, but never directly controlling it. I would not consider ADAPTASIA to be an improvised piece, but it still shares many qualities with improvisation, and has the potential to evoke some of the same effects.

One goal of ADAPTASIA is to serve as a simple entry point to the emotional expressivity of music improvisation. The music education field provides useful perspectives from which to view the benefits and limitations of amateur improvisation. The US National standards for Music Education has included improvisation since its inception, as do music education curricula in many European countries (Larsson and Georgii-Hemming 2019). Two broad categories are defined in the field: structured improvisation, which provides a clear framework and expects a student to follow Western music conventions to meet specific

outcomes, and free improvisation, which emphasizes open-ended creativity, exploration, and experimentation. While structured improvisation is usually implemented with an aim to develop musical skills and knowledge, free improvisation aims for students to learn about themselves and others through music-making, with potential benefits including improved empathy, innovation, flexibility, and social development (Ng 2021). However, free improvisation is relatively uncommon in music education, possibly due to a lack of comfort and experience with it among teachers (Larsson and Georgii-Hemming 2019). Although ADAPTASIA is in some ways highly structured, the user's interaction is explorative, personal, and free from expected outcomes, aligning it more closely with free improvisation.

What challenges might there be for a non-expert using improvisation for emotional expression? In his influential work "Growing with Improvisation," John Kratus (1991) outlines the factors that unify and distinguish expert improvisers from beginners, with an aim to focus pedagogy towards each developmental level. He defines three shared qualities among all improvisation: purposeful movements to produce sounds in time, the inability to revise the music after it has been played, and the freedom of the improviser to choose pitches and durations, sometimes within musical constraints. However, while beginner improvisation will focus primarily on explorative manipulation of the instrument, an expert must develop a variety of skills in audiation, musical structure, and stylistic conventions in order to produce more compelling music. Kratus also notes that while an expert improviser is able to match their musical intention to the musical result, a beginner may struggle to manipulate an instrument at a level where the sounds produced can fluidly match their intentions. This struggle persists to some degree until a person is a master of their instrument, after years of diligent practice and training that most people will never achieve. The development of musical intuition, too, takes

extensive time and effort, and can be a barrier for most people to reaching a high level of comfort and skill in improvisation.

The ADAPTASIA interface limits and simplifies the amount of control that the user has over the musical result. This has the benefit of increasing the accessibility of emotional expression through music for people who have limited technical musical skills, but also a potential to excessively stifle creativity and agency. Is there any benefit to using a tool to generate complex but unoriginal music, rather than expressing yourself directly with the musical abilities that you have? I believe the answer is determined by context and personal preference. Elementary educators have found social and emotional benefits to free improvisation among children with little musical experience (Ng 2021), and I know that my own experience of improvisation couldn't be matched by a tool which does most of the work for me. However, I have also found that the limits to my musical abilities can feel stifling themselves; after 14 years creating music, I still feel a frustrating disconnect between the music I want to express and the music I am able to express, and people without experience in improvisation - even expert musicians who have only ever played from sheet music - often find it intimidating and inaccessible. By abstracting musical intention to simple dials, I believe ADAPTASIA demonstrates a way to lower the barrier of entry into personal expression through music, providing emotional stimulation in a way that balances the interactivity of improvising and the passivity of listening.

The aleatoric - or random - elements in ADAPTASIA are also worth examining in relation to improvisation. My project uses aleatoric principles to imitate improvisation, even without any input from the user. Organizational factors - such as which musical fragment plays out of several options or whether a transition occurs - are guided by random numbers generated

by the computer. Many of these factors are analogous to choices that would be made by a performer improvising with an ensemble, but rather than being guided by any musical intuition, the computer acts only with cold randomness. Every sound is pre-recorded and organized in a specific way, and every musical result is derived from my stochastic design. The random elements in ADAPTASIA are integral to its ability to loop forever without appearing to repeat. At an increased scale, this type of musical design can produce a functionally endless variety of music. However, the lack of purpose behind the computer's random choices disqualifies ADAPTASIA from being considered "improvisation" and places it squarely in the realm of indeterminacy.

ADAPTASIA and Indeterminacy

Musical indeterminacy is most famously associated with John Cage, who shook the world of Western art music with pieces like *Music of Changes* (1951), composed by using a Chinese divination text called the *I Ching* to determine each musical element, and *Solo for Piano* (1958), consisting of 63 pages of abstract notation to be played "wholly or in part and in any sequence" (Cage 1958) Indeterminacy incorporates chance into a piece beyond the usual minor variations of interpretation between performances. The chance may be in respect to composition, as in *Music of Changes*, which contains the same randomly-originated notes and rhythms in each subsequent performance, or it may be in respect to performance, as in *Solo for Piano*, which may vary widely between performances based on the performer's choices. ADAPTASIA is indeterminate with respect to performance, so I will restrict my discussion to this type of indeterminacy.

Indeterminacy in Western art music has always been a heated topic in discussions of musical aesthetic and value. There was a sense in the 1950s that conventional Western music had been exhausted of its possibilities and that bold ventures away from tradition were needed in order to inject vitality into the field. Emerging alongside John Cage were integral serialists, who took a highly mathematical and deterministic approach to music by composing with numerical series representing pitches, rhythms, dynamics, and other musical aspects. Cage reacted in the opposite direction and took the stance that determinate music, fully and statically controlled by the composer, was untrue to the uniqueness of each moment in life and served as an excessive assertion of the composer's own ego (O'Grady 1981).

The most common criticisms of both indeterminacy and serialism were (and continue to be) an aesthetic distaste for the music, which could come across as grating and meaningless, and

a sense that the composer ought to have artistic control and responsibility over their work. My own reaction as a composer is to be protective of my own authorship over my music; I often put great effort into structuring and designing it to match my aesthetic vision, and when I listen to music, it is often those efforts which inform my enjoyment of the piece. Heavy indeterminacy or serialism in music can detract from the impressiveness, humanity, and simple beauty that makes it worth listening to for the average person. However, the ideas behind indeterminacy can be artistically useful even for a composer who values traditional enjoyability and "assertion of ego."

ADAPTASIA has a carefully predefined structure which incorporates randomness and user input for the purpose of variety and user interactivity. Its indeterminate qualities are more utilitarian and less philosophical than those typical of the mid-century rise of indeterminacy - it is intended not as a purely aesthetic innovation but as a functional permutation rooted in the tradition of video game soundtracks.

Case Study: Terry Riley, *In C* (1964)

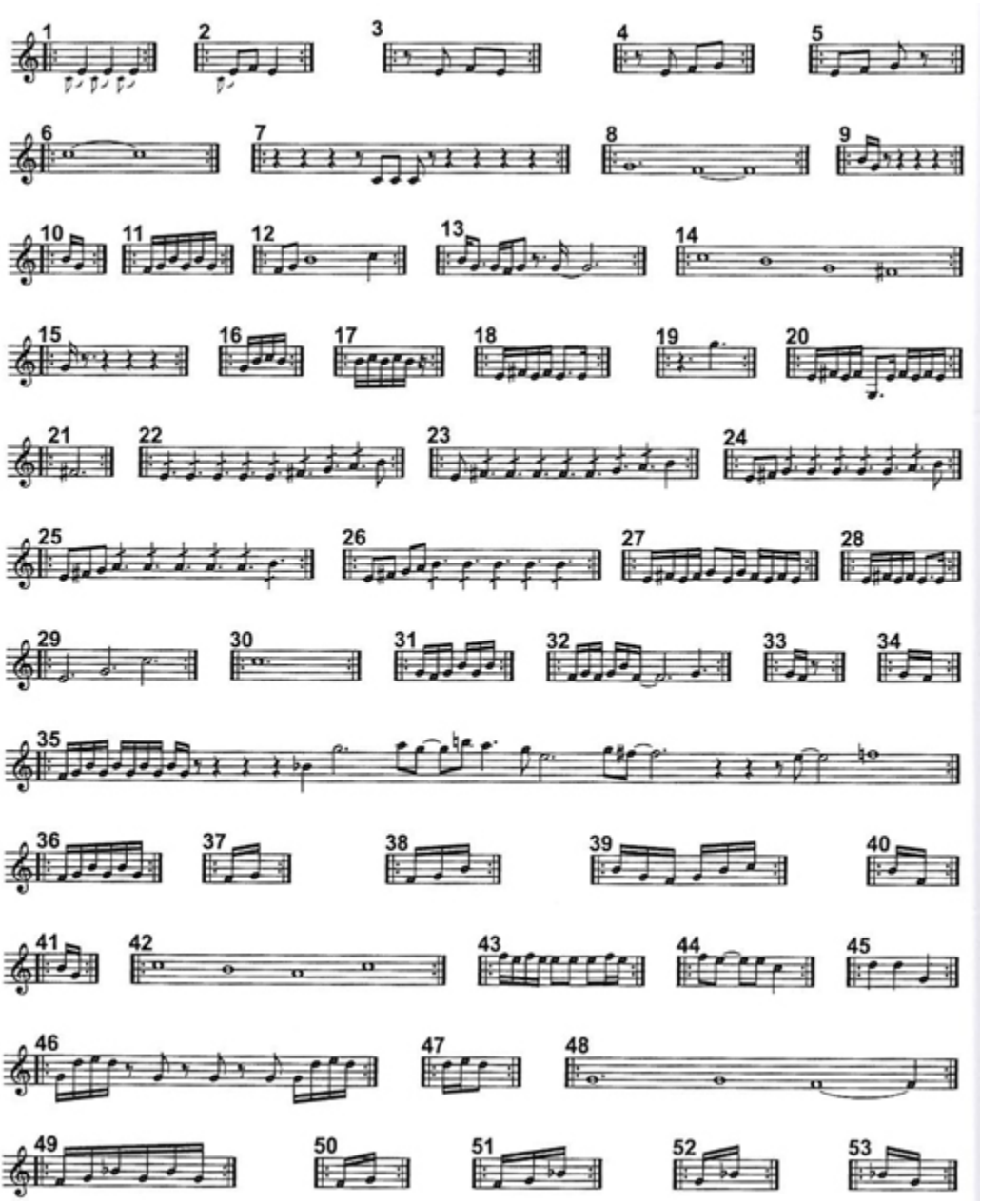

Figure 9: Score of *In C* (copyright Terry Riley, 1964)

An early and quintessential minimalist piece, *In C* consists of 53 short musical phrases which may be repeated an arbitrary number of times by each individual performer. As suggested by its title, the piece maintains a fundamental harmonic stasis, supported by the metronomic

pulse of C customarily played in repeated 8th notes on a piano. The phrases must be played in order, although any may be skipped, and the ensemble is instructed to try to keep within two to three phrases of each other.

Though *In C* is written for an indeterminate number of human players, the aleatoric method of sequencing short, pre-written modules of music is well-suited to computers. Robert Carl describes *In C* as "a piece of software … a series of rules and predefined relationships that execute a task; the user can then customize input and tweak aspects of the rules and relations to produce a product that is regarded as personal" (Carl 2009). ADAPTASIA, as an actual piece of software that sequences short segments of music, can be seen as an extension of this idea of software-like music. Like *In C*, automatic variation in ADAPTASIA is created by random 'choices' made by different 'performers': the selections of tracks within multi-instruments and the randomization of the structure. The human performer joins the ensemble to add in their own choices through the user-configurable parameters. The user of ADAPTASIA has an experience not unlike a single performer in *In C*, who contributes their arbitrary musical decisions on how many times to play each segment to the collection of decisions made by the rest of the ensemble within a highly structured framework. However, the user of ADAPTASIA has more control over overarching musical factors than a single *In C* performer would, as if the conductor of *In C* could direct the whole ensemble to change their performance.

Case Study: Brian Eno, *Thursday Afternoon* (1985)

Brian Eno's sparse, undulating tracks have defined the genre of "ambient music" ever since he popularized the term in the 1970s. In a *New York Times* article, John Pareles relates the psychological space of Eno's music to the contemplative fog of the COVID-19 pandemic: "Amorphous, open-ended, unstructured time, with undercurrents of foreboding, pockets of

boredom and fleeting interludes of peace or reassessment" (Pareles 2020). Eno's dreamy soundscapes have continued to captivate listeners decades after release, inducing meditative, nostalgic, and transformative experiences.

Thursday Afternoon released in 1985 with a single 61-minute track that had served as an audio accompaniment to an avant-garde "video painting" production. My first listen immediately sent me into a relaxing afternoon nap, and my second provided a peaceful background to my work despite my usual aversion to study music - the happenstance melodies over a thin, grainy synth drone were not organized enough to be distracting. The piece is constructed of several looping tracks, each playing a single note which repeats at a given time interval, over a constant synthesizer chord which slowly evolves throughout the piece. Because the intervals between note repetitions are different for each track (and slightly randomized within each track), the notes always line up differently, creating quasi-melodic fragments that are generated by the system of loops rather than deliberately placed by the composer. The generative technique of simultaneously looping tracks of different lengths allows for the composition to theoretically play endlessly without noticeable repetition.

I employ a comparable technique in one of the more ambient layers in ADAPTASIA: the 'static' and 'wind' event instruments. I recorded the individual notes of a B flat major 7 chord on a synthesizer, each note lasting about 19 seconds, and created two simultaneous multi-instruments that would each choose randomly among the four notes. The second multi-instrument has a delay offset randomized between 5 and 10 seconds, meaning that it will begin to play sometime between 5 and 10 seconds after the first multi-instrument begins to play. The sound is triggered at even intervals, so the effect is an endless drone that shifts among the notes of the chord at partially-randomized intervals.

Additionally, as a tape piece, *Thursday Afternoon* occupies a similar philosophical niche to ADAPTASIA with regard to its status as a "live performance." In the quotation at the beginning of this section, Paul Turowski describes a tape piece as a composition which "exist[s] as a 'fixed' medium" but "is still performed, even if somewhat automatically" (Turowski 2016). John Cage and some other 20th century musicians were famously opposed to recordings of music performances, considering them a false objectification of a naturally effervescent art form. This attitude would be futile today, where few would argue against the value of music recordings and significant quantities of highly regarded music only exist as computer "performances." However, tape pieces, a specialty of both Cage and Eno, turn recorded sound into a "performance" by organizing several tracks of sound to interact with each other in real time. ADAPTASIA does the same, placing it towards the "performance" end of the ambiguous recording-performance spectrum despite its fundamentally prerecorded nature.

Adaptive Music in Video Games

A soundtrack is most effective when it efficiently pulls focus to important information on screen. In a film, the visual elements have been edited into their final linear form, and the composer watches the completed film in order to write a seamless musical overlay - each moment can be perfectly timed in order to draw focus to the setting, characters, or narrative. Achieving this kind of precision becomes immensely more complicated when scoring for games, where core aspects of timing are under the control of the player. The task of improvising a soundtrack that dynamically adapts to the player's choices is one that seems suitable only for a highly intuitive musician playing along in real time. How can we possibly construct a precomposed soundtrack that can handle every possible sequence of events in a game?

One solution is to use a single looping track to highlight the general aesthetic of an area or game level, relying on sound effects and basic transitions between tracks to provide interactivity. This is a simple and overwhelmingly common approach: for instance, in the platformer *New Super Mario Bros* (2006), a looping music track establishes the atmosphere of each level, while sound effects call attention to enemies, platforming obstacles, and Mario's movement. There are only a few interactions between the player and the background music: the music stops if Mario dies or finishes the level, a special track plays while Mario is equipped with a star powerup, and the music increases in speed when there are 100 seconds left on the level timer. Similarly, in *Mario Kart 8* (2014), each racetrack has a signature music track which energizes the player and contributes to the atmosphere, while gameplay information is conveyed through sound effects as players navigate the course and throw, equip, or are hit by items. The music speeds up on the final lap to convey urgency, then stops when the race is finished.

However, the interactive medium of video games presents opportunities for more complex musical interactivity. Adaptive music, as it is most commonly known, aims to create seamless transitions between music tracks and allow the soundtrack itself to react to the player's actions. In the real-time strategy game *Pikmin 3* (2013), each piece in the soundtrack has several versions which play simultaneously, but fade in and out based on the game events. Similarly to the aforementioned games, time and progression is indicated by a sunset version of the music which fades in as the timer on each day begins to run out, but there is also a more urgent version to indicate the presence of an enemy, and a more percussive version during combat. Musical changes indicate other important information relevant to and dependent upon gameplay, such as when a task has been completed offscreen, by fading instruments in or out: each instrument gains an association with in-game information. Adaptive music is a vital tool in many games today for both capturing the spirit of the gameplay experience and aiding gameplay itself.

Video Game Analysis

Academic analysis of video games and their soundtracks is an emerging field with little direct lineage. Video games themselves have a short history of rapid evolution alongside the technology of the past half-century, and it may be a long time before the genre approaches the legacies of literature, theater, and film. Some scholars have begun by engaging with video games through previously established lenses such as film analysis - in many ways a fitting analogue to video games given its audio-visual nature and close relationship with technology. However, film analysis does not take into account the significant impact of interactivity on the design and experience of video games. Another technique, derived more from product testing and social science, isolates elements of video games to directly test their impact on players (Klimmt et al. 2019, Poppelaars et al. 2018); these studies take a concrete approach to analysis

which may sometimes have too limited a scope to be creatively relevant to game designers. There also exist educational resources and academic case studies on adaptive soundtracks in games, which shed light on the practical considerations of composing for games (Chance 2016, McAlpine 2016, Sporka and Jan 2017). However, the best resource on the current state of video games and their soundtracks remains the games themselves, so I will conclude with two case studies of recent independently developed games whose FMOD-implemented adaptive soundtracks contribute to their emotional interactivity (Matt Makes Games 2018, House House 2019).

Journals of aesthetics, culture, and musicology house a few perspectives on the artistic analysis of video games. Hart (2014) leans into film music analysis and phenomenological approaches to the study of video game music in his musicology paper, "Meaningful Play: Performativity, Interactivity and Semiotics in Video Game Music." He situates the player's interactions as performative acts which interpret the musical experience envisioned by the composer just as they interpret the game experience envisioned by the game designers.

Grimwood (2018) focuses on cultural genealogy in "Procedural monsters: rhetoric, commonplace and 'heroic madness' in video games," as he examines how video games use shared rhetoric of madness such as the monstrous double and the reaching tentacle to make sense of mental health experiences. The author claims that the cultural genealogy of madness is vital for our understanding of fictional portrayals of mental illness, and that even tropes that historically have been used in a shallow and othering way towards mentally ill people can be "reobjectified" and reflexively engaged with in modern media, especially when the "heroic" player character is the one exhibiting the tropes (Grimwood 2018). Video games are built on a fundamental control over the environment through button inputs, which can be effectively

juxtaposed with the loss of control that is associated with madness. In the specific tropes examined by the author, this is seen in the uncontrollability of the monstrous double and the reaching tentacle, despite their attachment to the player character.

Building on similar ideas of identity, Antonsen (2021) takes a linguistic view of game interactivity in "Self-Location in Interactive Fiction." This paper theorizes on the nature of a player's projection onto a player character in a video game and differentiates it from self location in non-interactive media. The author makes a linguistic case for the player truly imagining that they are someone distinct from themselves, noting that a person playing a game will often verbalize the player character's actions as their own (I explored the area, I killed the enemy, etc). The author then analyzes the separation between the player's self and the self they inhabit in the player character, claiming that the player keeps track of two separate identities simultaneously with separate motivations and processes of decision making (Antonsen 2021). I think that some of this analysis may be misplaced or up for reinterpretation - for instance, speaking about actions in a game with "I" statements may be more of a fact of interactivity than a reflection of identity projection. While a person watching a movie doesn't do much but observe and eat popcorn, a person playing a game is performing actions which can be more usefully described as "I defeated the boss" than "I pressed a sequence of buttons on my controller."

All three of these papers first view video games from a lens applicable to non-interactive media, then examine the impact of interactivity. This separation highlights the lack of analytical foundation we currently have for interactive media - rather than using an established academic familiarity with the genre to analyze video games as a whole, there is a tendency to ground interpretations in an adjacent genre before probing the further considerations that need to be made to account for the player's interactions with the game. Such grounding is useful for the

relations it draws to more firmly established concepts in film and literature analysis, but it is also a clear signifier of the limitations and newness of video game analysis. Just as video game composition is limited by dependence on techniques for film composition, where each musical feature can be timed to a consistent visual timeline, video game analysis is limited by its dependence on a legacy of analysis of non-interactive media.

The limitations of comparisons between interactive and non-interactive media are clearly visible in a 2019 study of the effects of soundtrack music on the cognitive and affective video game experience (Klimmt et al.) which does not examine the interactivity of the music - only its presence or absence. Based on previous research on the effects of film soundtrack music, the authors hypothesized three effects of congruent game soundtrack music: eliciting and amplifying emotions, intensifying spatial presence (the feeling of being physically located in a virtual environment), and intensifying identification with the player character. Conducting an experiment where participants played games with and without their soundtrack audio, they observed increased enjoyment related to emotions elicited by the soundtracks. However, while they did observe increased spatial awareness, they did not find an increased sense of identification or see increased enjoyment related to either of these factors, thus not finding the cognitive route that they expected from film soundtrack research. The limited consideration of the interactivity of the music mirrors similar studies of film music, but ignores the differences between film and video games that may be the most relevant to soundtrack experience.

In a similarly limited study from a game design perspective, Poppelaars et al. (2018) tested the effects of explicit mental health messaging in video games on motivation and affective experience of players. When participants played a game after viewing a trailer emphasizing either its mental health benefit or its entertainment value, the researchers did not observe any

differences in intrinsic motivation or competence between the groups, but did observe less reports of autonomy in the mental health group, indicating that the group who saw the game as a mental health aid felt limitations on their freedom. The ramifications of mental health messaging are relevant for my game and others with relations to mental health, and relate to the artistic exploration of mental health rhetoric in games by Grimwood (2018). However, rather than using a game with mental health themes for their research, the team chose a game with no relation to mental health, claiming that no such game existed that could also pass as an entertainment-focused game. While this may be true when examining the limited catalog of games intended for clinical psychological use, the researchers missed an opportunity for a more convincing and artistically relevant study with a game like *Celeste* (2018) or *Night in the Woods* (2017), which combine legitimate entertainment value with mental health themes.

Adaptive Soundtracks

In his instructional book *Composing Music for Games* (2016), Thomas Chance breaks down techniques for composing adaptive soundtracks into two categories: vertical reorchestration and horizontal resequencing. Vertical reorchestration in the technique of fading in different simultaneous tracks of music based on game parameters, such as location, danger, progression, or any possible aspect about the state of the game that you might want to associate with a musical change. This is also a common and flexible game music technique which allows for a lot of specificity and granularity - each instrument in a piece might have a different relationship to the game, and the combination of instruments can then house a highly specific set of emotional/logistical information for the player. Horizontal resequencing is the technique of moving to a different point in the timeline of a piece based on game parameters - for example,

looping certain regions of a piece until a certain parameter is set by the game code, at which point a transition is triggered into a different part of the piece. This allows the music to progress according to game events, and presents opportunities for transitions to be written beforehand to cover any combination of game events while still sounding precomposed.

A few papers exist which cover case studies of adaptive music implementation. McAlpine (2016) examines a proof-of-concept gestural music interface called 'BitBox!' designed to teach interactive music composition. The interface deconstructs the technical and creative frameworks of adaptive video game music and facilitates interactive exploration of adaptive music techniques. The author claims that most university students studying creative sound production struggle to apply their composition skills to real-time adaptive music because they lack an understanding of the necessary concepts and approaches. In the Bitbox! interface, a button triggers an event-driven music cue and several ultrasonic sensors allow for resequencing, reorchestration, and generation based on the proximity of the user's hand to the sensor (similar to a Theremin). Users found the system natural, expressive, and learnable.

Sporka and Jan (2017) showcase a piece of adaptive music middleware created for the video game *Kingdom Come: Deliverance*. The authors wanted to make transitions between different tracks of orchestral music sound as seamless as if they had been composed directly using a network of endings and intros, as well as rising and decaying sound effects called cinels to disguise crossfades for near-instant transitions into combat music. The related work section discusses data sonification, vertical and horizontal music layering techniques, and algorithmic music (such as aleatoric, generative, and neural network based compositions). The authors describe its implementation in *Kingdom Come* as a case study with useful descriptions of the

type and organization of tracks, and present a user study indicating that players could not differentiate between an adaptively generated transition and a composed transition.

Case Study: *Celeste* (2018)

Celeste is a video game where you play as a girl named Madeline who is attempting to climb to the top of Celeste Mountain. The game translates experiences of depression and anxiety into interactive gameplay and audio systems. For instance, the game features a minigame where the player repeatedly holds and releases a button to mimic blowing a feather into the air, causing the music to become more calm as Madeline recovers from a panic attack. The villain of *Celeste* is a shadow-version of Madeline (referred to by the developers as 'Badeline') representing the mean and insecure parts of herself that she wants to be rid of. After spending most of the game fleeing Badeline, Madeline learns that by accepting and comforting this part of herself, they can become more powerful (represented by an extra dash, enhancing the platforming mechanics) and reach the summit of the mountain together. The team behind the *Celeste* soundtrack published the full FMOD project used for the game for free access, making it possible to view the mechanical construction of each adaptive musical track and sound effect (Power Up Audio, 2018).

Case Study: *Untitled Goose Game* (2019)

Untitled Goose Game is an indie video game in which you play as a goose with a list of mildly nefarious tasks to complete around a quiet English village. The tasks often involve startling humans to cause minor injury or stealing items and moving them to amusing spots. The game implements an FMOD-powered soundtrack with high- and low-tension versions of

Debussy's 12th Prelude which underscore the stealth and chase sequences that arise during the mischief-making gameplay. The adaptive piano music in the game was inspired by people's misunderstanding of an early trailer for the game, where visuals were cut around the music but appeared to be influencing the music. In order to make this misconception a reality, the team recorded a high energy and low energy version of each piece and split them into two-beat chunks which are handled by the music logic in the code. Depending on the game action, the music system plays the next chunk in either its high energy form, its low energy form, or waits in silence ("Interview: The sound of *Untitled Goose Game*"). The highly modular approach to horizontal resequencing in *Untitled Goose Game* allows it to uniquely adapt to gameplay situations, reflecting the player's actions in the excitement or relaxation of the music almost immediately.

Reflection

This project was subject to several changes in scope and focus, some of which can highlight the limitations and surprising complexity of building highly modular music systems. The background research I did, guided by the connections that my thesis committee and I drew between video game soundtracks and other areas of music history, performance, and education, also led me to wonder about the musical agency commanded by myself as the 'composer' of ADAPTASIA as compared to its user, and as compared to the computer itself. Here, I reflect on the role of musical intention and agency in adaptive music, as well as which of my design choices may be worth reconsidering in the future.

Who is the performer of ADAPTASIA? Who is the composer? Who holds musical agency in an adaptive piece, where the composer's decisions combine with external factors and performer decisions to create the end result? Who holds musical agency in a performance of a Beethoven symphony, where the composer's decisions lay out a strict sequence of actions for each member of the orchestra, every single performance? Is a person holding a radio a performer of adaptive music? What about a person selecting an auto-generated 'Sad girl vibes' playlist on Spotify? Adaptive music can expand and combine the ordinary roles of composers and performers.

How much musical intention should be put into a complex adaptive piece? I found myself attempting to design ADAPTASIA perfectly so that I could be confident that any possible outcome would sound 'right'. Did people like John Cage and his contemporaries care about that? It's easier to perceive their designs as flippant, and a less careful design would indeed drastically reduce the design burden for an indeterminate piece, where the number of possible outcomes is so high compared to the one option in a standard piece.

A music score usually shows one predetermined sequence of events, and the composer designs that one sequence according to their liking. This is enough of a challenge to achieve that most composers never stray from this linear design, just as most books and movies do not allow their audience to choose among different options for how the story might go. And just as allowing choice in a book or movie makes it immediately much more difficult to ensure that all possibilities are well-designed and satisfying, allowing choice in music multiplies the burden on a composer who wants to maintain creative control over the whole piece. But is it really necessary to maintain control? Taking a more hands-off approach makes indeterminate composition more viable by shifting the artistic agency of the piece and reducing its burden on the composer.

Let us consider the effect that the range of adaptive possibilities may have on the logistical viability of adaptive music, especially for video games: some of my efforts to make ADAPTASIA highly modular and non-repeating may not have resulted in an experience that has significant benefits over a simpler structure. To build the sets of music clips that made up a single two-bar multi-instrument in ADAPTASIA, I recorded myself improvising on my keyboard for several minutes using a Digital Audio Workstation (DAW) called Reaper, then edited out rhythmic errors, selected a set of the most successful takes, and exported them into individual mp3 files. Then, using an audio editing program called Audacity, I faded out the ends of each mp3 file so that the sound wouldn't cut out too suddenly. Finally, I moved the edited files into FMOD Studio and dragged them into position.

In many cases, such as the Piano+ and Piano- tracks, each music clip is almost identical, differentiated only by the minor variations between performances of the same music. My intention was to more closely approximate the experience of live music by recording my live

performances, which contain humanizing nuances in individual note volumes and rhythms, and randomly cycling among several similar performances in FMOD so that these nuances don't become stale through over-repetition. However, when considering how much more easily a similar experience could have been created with longer clips, fewer takes, or computer-generated performances, I think that consideration should be given in future similar projects to which efforts are truly worthwhile. Is non-repetition even a worthwhile goal, or does it detract from the catchiness that makes a game soundtrack enjoyable, or the predictability that makes adaptive variations a useful signal of information? It would be worth performing user experience studies which consider the tangible impact of these variables in order to optimize audio-production efforts when time is limited.

Most indeterminate pieces which redirect agency to the performers require that the performer have enough musical skill and intuition to be able to create meaningful music from instruments given vague instructions. Jazz musicians train extensively to develop their improvisation skills, which are built on technical proficiency, music theory knowledge, listening to lots of jazz music, and practicing extensively to develop intuition and style. Improvising music can be a meaningful and cathartic experience, but it is restrained by the musical skills of the performer. ADAPTASIA simplifies the agency of its performer into an abstract set of controls that limit the possible outcomes to a narrow range that I define. In further experiments with adaptive music, I would like to explore the benefits and drawbacks of widening this range, giving the performer more of an opportunity to exert their own musical intentions while relaxing my grip on my own.

Reference List

Antonsen, Paal Fjeldvig. "Self-Location in Interactive Fiction.". *British Journal of Aesthetics* 61, no.1 (2021): 41 - 52.

Cage, John. 1951. *Music of Changes*.

Cage, John. 1958. *Solo for Piano*.

Carl, Robert. 2009. *Terry Riley's In C*. Studies in Musical Genesis, Structure, and Interpretation. Oxford ; New York: Oxford University Press.

Clements-Cortés, Amy, and May Tianyang Yu. 2021. "The Mental Health Benefits of Improvisational Music Therapy for Young Adults." Canadian Music Educator / Musicien Educateur Au Canada 62 (3): 30–33.

Cormier, Ramona. 1975. "Indeterminancy and Aesthetic Theory." Journal of Aesthetics & Art Criticism 33 (3): 285. doi:10.2307/428354.

Cox, Christoph, and Daniel Warner, eds. *Audio Culture: Readings in Modern Music*. New York: Continuum, 2004.

Eno, Brian. 1985. *Thursday Afternoon*.

Grimwood, Tom. "Procedural monsters: rhetoric, commonplace and 'heroic madness' in video games.". *Journal for Cultural Research* 22, no.3 (2018): 310 - 324.

Hart, Iain. "Meaningful Play: Performativity, Interactivity and Semiotics in Video Game Music.". *Musicology Australia* 36, no.2 (2014): 273 - 290.

House House. *Untitled Goose Game*. House House. Switch. 2019.

"Interview: The sound of Untitled Goose Game." Accessed October 20, 2021. https://www.fmod.com/blog/untitled-goose-game-interview.

Klimmt, Christoph, Daniel, Possler, Nicolas, May, Hendrik, Auge, Louisa, Wanjek, and Anna-Lena, Wolf. "Effects of soundtrack music on the video game experience.". *Media Psychology* 22, no.5 (2019): 689 - 713.

Kratus, John. “Growing with Improvisation.” *Music Educators Journal* 78, no. 4 (December 1991): 36–40. https://doi.org/10.2307/3398335.

Larsson, Christina, and Eva Georgii-Hemming. “Improvisation in General Music Education – a Literature Review.” *British Journal of Music Education* 36, no. 1 (March 2019): 49–67. https://doi.org/10.1017/S026505171800013X.

Matt Makes Games. *Celeste*. Matt Makes Games. Switch. 2018.

McAlpine, Kenneth. "BitBox!: A case study interface for teaching real-time adaptive music composition for video games.". *Journal of Music, Technology & Education* 9, no.2 (2016): 191 - 208.

McPherson, Malinda J., Monica Lopez-Gonzalez, Summer K. Rankin, and Charles J. Limb. “The Role of Emotion in Musical Improvisation: An Analysis of Structural Features.” Edited by Howard Nusbaum. PLoS ONE 9, no. 8 (August 21, 2014): e105144. https://doi.org/10.1371/journal.pone.0105144.

Ng, Hoon Hong. “The Value of Learning Collective Free Music Improvisation: Preservice Music Educators’ Perspectives.” *Journal of Music Teacher Education* 30, no. 3 (June 2021): 40–53. https://doi.org/10.1177/10570837211018287.

O’Grady, Terence J. “Aesthetic Value in Indeterminate Music.” *The Musical Quarterly* LXVII, no. 3 (1981): 366–81. https://doi.org/10.1093/mq/LXVII.3.366.

Pareles, Jon. 2020. “Brian Eno’s 15 Essential Ambient Works.” *The New York Times*, April 29, 2020, sec. Arts. https://www.nytimes.com/2020/04/29/arts/music/brian-eno-ambient-songs.html.

Poppelaars, Marlou, Anna, Lichtwarck-Aschoff, Marloes, Kleinjan, and Isabela, Granic. "The impact of explicit mental health messages in video games on players’ motivation and affect.". *Computers in Human Behavior* 83 (2018): 16 - 23.

Power Up Audio. *Celeste FMOD Project and Source Audio Content*. V. 1.00.00. Firelight Technologies Pty Ltd. FMOD Studio. 2018.

Riley, Terry. 1964. *In C*.

Sporka, Adam J., and Jan, Valta. "Design and implementation of a non-linear symphonic soundtrack of a video game.". *New Review of Hypermedia & Multimedia* 23, no.4 (2017): 229 - 246.

Thomas, Chance. *Composing Music for Games: The Art, Technology and Business of Video Game Scoring*. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2016.

Thomas, Philip. "Understanding Indeterminate Music through Performance: Cage's *Solo for Piano." Twentieth-Century Music* 10, no. 1 (March 2013): 91–113. https://doi.org/10.1017/S1478572212000424.

Turowski, Paul. "Digital Game as Musical Notation." Dissertation, University of Virginia, 2016.

"Unity." 2020. Unity Technologies. https://unity.com/.