Procedural Rhythm Generation

for the Hierarchical Wave Function Collapse Model

Ágnes Mikó[1]
Supervisor: Dr. Rafael Bidarra[1]
[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 28, 2024

Name of the student: Ágnes Mikó
Final project course: CSE3000 Research Project
Thesis committee: Dr. Rafael Bidarra, Dr. Joana de Pinho Gonçalves

Procedural Rhythm Generation

for the Hierarchical Wave Function Collapse Model

Ágnes Mikó

Delft University of Technology
Delft, Netherlands

ABSTRACT

Music, a profound expression of the human experience, is built upon intricate patterns that humans have spent centuries trying to understand and recreate. Consequentially, Hierarchical Wave Function Collapse was proposed as a novel model for procedural music generation, which organizes musical constraints within a three-level hierarchy. However, music with melody but no rhythm, is like a text with all characters but no spacing. This research aims to create rhythm procedurally by exploring existing representations and combining the most effective elements to propose a simple but comprehensive rhythmic model. Though not all-powerful, the proposed model accommodates varied measures and motifs. By implementing it, the research evaluates its efficiency and limitations, identifying the types of music it cannot generate. This study aims to contribute a tool for mixed-initiative music creation, expanding procedural music generation with a nuanced approach to rhythm.

CCS CONCEPTS

• **Applied computing → Sound and music computing**.

KEYWORDS

wave function collapse, procedural music generation, constraint programming, rhythm generation

ACM Reference Format:
Ágnes Mikó. 2024. Procedural Rhythm Generation: for the Hierarchical Wave Function Collapse Model. In *Proceedings of Research Project (CSE3000)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Music stands as the heartbeat of the human experience, capturing the profound beauty of existence [10]. The definition of music being "any rhythmically repeated occurrence" even allows us to compare it to the music of whale and bird species [5]. While compressing its richness into one definition is absurd [2], music can be investigated from various perspectives including tonal scales, repetitive structures of pitch, and the periodicity of beats. This paper explores the latter, delving into the definition and composition of rhythm.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM
https://doi.org/XXXXXXX.XXXXXXX

Recent advancements have introduced Hierarchical Wave Function Collapse as a novel approach to procedural music generation [16]. This method enables the creation of melodic structures by setting various constraints organized in a three-level hierarchy. As a proof of concept, a music generator was implemented using this method [15]. However, as music cannot exist without rhythm, this paper proposes a generative model to help guide rhythm creation within Hierarchical Wave Function Collapse.

This brings us to a crucial question: How can we integrate rhythmic patterns into the Hierarchical Wave Function Collapse model? Several related thoughts emerge:

(1) What exactly is rhythm, and what essential constraints define it?
(2) How have other music generators approached the modeling of rhythm?
(3) What techniques can enable more flexible rhythm creation?
(4) Considering its application as a tool, what would be an intuitive visual representation of rhythm?

Hence, the objective of the research is to explore the definition of rhythm and its existing models and present a novel rhythmic model that serves as a tool for mixed-initiative music creation, not only in general terms but specifically tailored to the style of HWFC. The goal is not to build a music generator that can replace human creativity, but one that can produce varied outputs that conform to given constraints, serving as possible inspiration for artists.

The paper provides an overview of the research background and the work conducted throughout the study, detailing the inner workings of the resulting model. It also analyzes the outcomes of the model and addresses the ethical considerations involved. Finally, the paper concludes with a discussion that contextualizes the findings within a broader framework, highlighting the inherent but resolvable limitations.

2 PREVIOUS WORK

This section presents Hierarchical Wave Function Collapse (HWFC) and explains how it is used in procedural music generation. It also discusses approaches of previous works to defining and representing rhythm, including an examination of rhythmic semantics and procedural content generation techniques. This exploration shows how different models and representations can enhance the complexity of rhythm in HWFC-generated music.

2.1 Wave Function Collapse

Wave Function Collapse (WFC) is a constraint-solving algorithm developed by Maxim Gumin [8] inspired by quantum mechanics, which generates structured content by iteratively collapsing a grid of cells to definite states based on input sample constraints, similar to the quantum superposition. Karth and Smith discuss the algorithm's inner workings, its help in texture synthesis, and its various applications in the wilderness of procedural content generation (PCG) [6]. In broader terms, PCG refers to the creation of content (e.g. textures, music, rules, and quests in the world of video games) with the help of an algorithm [14].

Hierarchical Wave Function Collapse

Hierarchical Wave Function Collapse (HWFC) enhances this by introducing a hierarchy of layers called canvases, where each canvas inherits constraints from the one above it. The layers each represent their canvas with cells set to be collapsed based on the given constraints. As seen in Figure 1, these three layers comprise the music generation system:

- **Sections**: At the top level, the overarching structure of a piece (e.g., intro, verse, chorus) is defined.
- **Chords**: Each section contains a canvas of chords, representing the harmonic foundation. Here, chords are mapped to measures.
- **Melody**: Within each chord, a canvas of notes forms the melody, with pitches and durations specified.

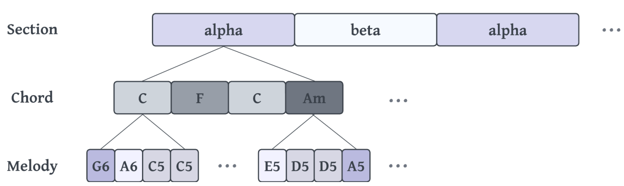

Figure 1: The layers of HWFC [16].

To demonstrate the conceptual workings of the generator, the system ProceduraLiszt [15] has been created. Upon closer look, it becomes apparent that a crucial part of the mixed-initiative process is the setting of constraints, which allows the artist the freedom to tweak the result to their liking, and to provide transparency about the inner workings of the model.

The rhythmic model applied by the system is quite simple. The parameters of minimum and maximum note length, pause length, and whether measures can start with a pause are used to create variations in the rhythm. However, this model lacks the sophistication needed to create more specific beats, limiting its flexibility in generating music.

2.2 The representation of rhythm

Rhythm can be interpreted in various ways, from the cadence of language to the changing of seasons and the beat of our hearts [3]. Often perceived as a periodic temporal shape of movement, rhythm defies straight substitution with concepts like pattern, order, or measure. Haili You argues that it encompasses a sense of suspense and fulfillment born from anticipation, a tension inherent in our perception of periodicity [17]. For the sake of simplicity, however, we will set aside the intangible Gestalt[1], and confine rhythm to a simplified generative model.

[1]Something that as a whole is more than the total of its parts.

There are a handful of recurring musical terms that appear throughout the paper and therefore have earned the merit of clarification. Schulkind [13] describes **the beat** as the simplest form of temporal structure. Beats, when appearing at regular intervals, create a pulse that makes music easy to follow. Some beats can carry more weight, creating a group of accented and unaccented beats, known as **the meter**. Schulkind points out the structural intonations imposed by meter, with longer notes typically starting on stronger beats, and shorter notes starting on weaker ones. In Western musical notation, this is defined by the **time signature**, which specifies the number of note values within a measure and is constructed by a lower numeral (indicating the note value or beat duration that is being counted) and an upper numeral (indicating the number of such note values in a measure). On the whole, Schulkind defines rhythm as "the serial pattern of variable note durations in a melody".

According to Rohrmeier, music is inseparable from a metrical interpretation [11]. He develops a model employing upbeat and downbeat notes to encompass syncopation, anticipation, and delay - elements that manipulate the timing of events, disrupting expectations of their positions. This model takes the form of a context-free grammar, defining each event as a triple $[a : b : c]$, representing a downbeat segment of duration b, an initial offset of a, and a final offset of c. The model addresses the rhythmic complexities mentioned by applying different rules to these components, like splitting or shifting, and as a result, highlighting the importance of less ordinary rhythms.

Several music generation systems use statistical models. Ames [1] discusses the usage of Markov chains and the potential meaning of the states within this framework. These states can be simple notes represented by pitch or duration, a vector of attributes, chords, or polyphonic phrases, or a moment representing all of the musical activity at a given time in the piece.

Machine learning has also proven useful in producing music. Loughran [7] reviews the use of genetic programming and evolutionary computation in music creation, with the conclusion that musicality and creativity are not easily definable concepts, nor is the fitness of a piece of music. The evaluation of the musical results appears to be almost futile in a way; rather, the emphasis should be placed on the availability of tools that facilitate its creation. This reinstates the importance of mixed-initiative methods in music generation.

Pearsall [9] presents a durational set theory model, representing rhythm as a set of proportions, each standing for the duration value of the note. The main advantage of this model is that it portrays equivalence between both ordered and unordered sets of durations.

MIDI editors represent music in a piano roll layout, where music is displayed as a function of time and pitch, as Figure 2 shows. This is similar to modern staff notation [4], with the difference being that note values are not indicated by their actual duration but by a corresponding symbol.

The simplest and most reliable representation among those discussed appears to be one based on the principles of Western tonal music, as it aligns closely with the MIDI output format used by ProceduraLiszt. When combined with the visual editing methods of MIDI editors, this could serve as a robust mixed-initiative environment for rhythm generation. Rhythm is a playful force of motion, and constraining it to the structure of WFC might limit

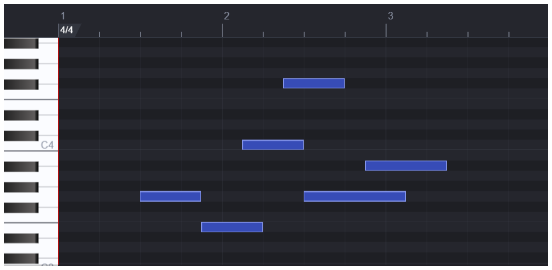

Figure 2: Music representation in a MIDI editor [12].

its expressive potential. The scope of this research is to explore the procedural generation of rhythm, not necessarily to adhere to the WFC algorithm. While WFC could be adapted for the model proposed in Section 3, the goal was to prioritize a method that allows for creative freedom and user interaction. This flexibility is essential for capturing the true essence of rhythm and aligning it with the mixed-initiative approach of ProceduraLiszt.

3 PROPOSED MODEL

This section gives an overview of a rhythm generation model inspired by the previous works in Section 2. As discussed earlier, there are many ways to generate rhythm, ranging from machine learning models to context-free grammars, to HWFC itself. The goal of this research is to create a simple rhythmic model that conforms to the style of HWFC. Emphasizing the use of constraints, and the focus on the mixed-initiative process, the model mirrors the principles of HWFC. While it does not directly use WFC for generating rhythm, it maintains the spirit of HWFC in its structure and user interaction, which the following subsections will expand upon.

3.1 A mixed-initiative approach

The scope of HWFC in procedural music generation was to be a tool for the mixed-initiative process. In the context of rhythm generation, features that support this process should allow control over the constraints and their factor of influence in the generation process. In an ideal situation, rhythm could also be created independently from these constraints, and it would be the system's responsibility to take that rhythm as a base, aligning it with the input constraints of time signature, meter, etc.

An important term in HWFC, a prototype is defined as "a template which acts as a meta-tile that can be applied on a slot of a canvas". Rhythm functions similarly to these prototypes, representing properties and constraints regarding the timing and duration of notes, that each repeating measure builds upon to create a consistent beat structure. In practice, to create the desired mixed-initiative environment, the rhythm could be manually defined by the user, in the manner of MIDI editors, and would draw over the underlying prototype.

3.2 Rhythm generation

To build the rhythm model, some decisions need to be made, specifically about: the definition of rhythm, a representation that correctly captures this, and a strategy to generate rhythmic sequences.

The definition. For the purpose of this paper, we augment Schulkind's rationale and define rhythm more precisely as "a pattern of variable-length duration notes and rests, summing up to the value determined by the time signature, loosely connected to the underlying beats of the meter". Heavy inspiration is taken from the modern staff notation, commonly used in Western tonal music. This system has stood the test of time, proving to be a reliable direction to follow.

The representation. In Western tonal music, two key attributes describe a note's rhythmic aspect: whether it is a note or a rest, and its duration. And so, generating rhythm involves combining sequences of different length note values and types. The resulting pattern will depend on the following constraints:

- the time signature
- the meter - with strong beats evenly spaced throughout
- the weight of the meter - the probability of longer notes appearing on strong beats
- the note values - a pool of values to generate from
- the melody length - the number of notes in a measure

The strategy. The process begins by generating an array of note lengths from a pool of possible values that add up to the measure duration set by the time signature. Longer notes are more likely to appear on the strong beats, depending on the meter. Each note is then randomly decided to be a rest or not, with an even distribution. This creates a rhythm for one measure, which is then combined with the melody of one chord to create a complete measure. Figure 3 shows an example of a possible rhythm.

Figure 3: A possible measure. The dark grey dots show the strong beats, and the light grey dots show the weak beats.

The resulting rhythm will be the pattern each measure of the piece builds upon. It will repeat throughout the piece, creating the strong pulse many have defined as the heart of rhythm.

In line with ProceduraLiszt's mixed-initiative process, users can either manually set a specific rhythm, or rely on the fallback generation strategy based on the predefined constraints. Within a measure, users can select note positions, durations, and whether they are notes or rests. If the measure has empty spaces, the fallback algorithm fills in the rest of the rhythm, which can vary over each measure. This will lead to a cyclical process illustrated in Figure 4 of adjusting the constraints while listening to the output, to achieve the desired piece of music.

4 IMPLEMENTATION & RESULTS

The visual editor allows users to choose a time signature, and a corresponding rhythm strip of one measure will be displayed. Lighter and darker semi-transparent alternating background colors indicate the beats. Users can drag and select where notes or rests appear within the measure. The algorithm has many constraints that can be adjusted to affect the rhythm generation. These constraints include the smallest unit of time displayed in the editor,

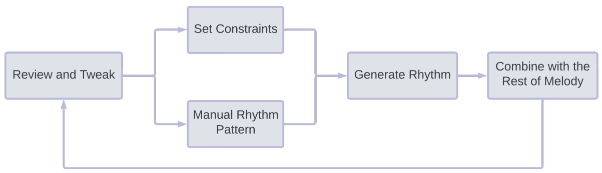

Figure 4: The process of rhythm generation.

the length of the melody, the possible note values appearing, the strong beats in ambiguous meters (e.g., 7/4 time signature), and the probability of longer notes appearing on strong beats. Figure 5 shows a measure in 5/4 time that begins with a quarter note, followed by a quarter rest, then a dotted eight note, the rest of the measure left empty to be filled out by the underlying algorithm.

Figure 5: Example of a 5/4 measure in the visual rhythm editor, with user-selected notes and rests.

Figure 6 shows the output produced by the previously set constraints. The four bottom lines represent the chords and indicate measure changes, aligning with each chord. Above these lines is the melody. As shown, each measure follows the predefined rhythm template, with the algorithm generating the end of each measure.

Figure 6: The generated output.

While the wide range of adjustable constraints and the ability to compose rhythm in the visual editor create a significant expressive range, there are certain types of rhythms that the model does not cover by its nature. These limitations will be discussed in the following section.

5 DISCUSSION

Within the model, rhythm is generated independently from the melody. Yet, in reality, this is rarely the case. Accentuation can be expressed in various ways, from holding a note longer to adding a slightly unexpected off-key tone. Both axes of rhythm and melody are meant to express, therefore they are bound to be connected, intentionally or not, their sum meaning more than either part alone. The model could benefit from a cross-relation between the two layers.

As previously mentioned, Rohrmeier noted the importance of less ordinary rhythms. The model does not cover cases of syncopation, delays, or polyrhythms. These phenomena all introduce value and uniqueness to rhythm, therefore it would be beneficial for the model to consider them.

As discussed in Section 2.2, the decision was made not to use WFC as the generation method. Nonetheless, since the generation still relies on constraints, the algorithm could be adapted to handle note values as durations in superposition and determine which cells to collapse according to the constraints, the concept of meter and strong/weak beats potentially being meta tiles.

6 RESPONSIBLE RESEARCH

Reproducibility is an important quality of research. It provides authenticity to the findings, by providing the steps to repeat the process and allowing to arrive at the same result. This is maintained in the current research, as following the described algorithm and implementing the model as it was detailed will yield outputs of the same expressive range.

Transparency is an essential quality for content generation in mixed-initiative art. It assures the logic of the process is clear to the user, and so establishes the necessary trust between the person and the tool. Fortunately, procedural music generation in HWFC is guided by setting constraints visible to and adjustable by the user. That assures that the parameters steering the generation process are clear to the user.

The generation of rhythm relies on no prior model training, as it is created solely based on input constraints, discarding concerns about data privacy. It is also important to note that the system serves as a tool to ignite inspiration, not to eliminate human creativity. Therefore, the intention is not to have great expressive power but to provide diverse examples that abide by the set of constraints.

7 CONCLUSION

The exploration of integrating rhythmic patterns into the Hierarchical Wave Function Collapse model has led us through a series of considerations. It began with defining rhythm and identifying its essential constraints. We proposed an algorithm for procedural rhythm generation, designed to match the spirit of the existing procedural music generator ProceduraLiszt. By implementing a rhythm generation model led by user-defined constraints taken from Western tonal music, we struck a balance between the algorithm and user control. The implemented model has demonstrated a humble but productive expressive range. Overall, this research explored what rhythm could mean within the context of procedural music generation and that of HWFC, improving an existing tool for mixed-initiative music.

REFERENCES

[1] Charles Ames. 1989. The Markov Process as a Compositional Model: A Survey and Tutorial. *Leonardo* 22, 2 (1989), 175–187. https://doi.org/10.2307/1575226

[2] Stephen Davies. 2012. On Defining Music. *The Monist* 95, 4 (2012), 535–555. https://doi.org/10.5840/monist201295427

[3] F. W. Flattely. 1920. Rhythm in Nature. *Science Progress in the Twentieth Century (1919-1933)* 14, 55 (1920), 418–426.

[4] T. Gerou and L. Lusk. 1996. *Essential Dictionary of Music Notation: The Most Practical and Concise Source for Music Notation.* Alfred Publishing. https://books.google.nl/books?id=MqAzAQAAIAAJ

[5] Patricia M. Gray, Bernie Krause, Jelle Atema, Roger Payne, Carol Krumhansl, and Luis Baptista. 2001. The Music of Nature and the Nature of Music. *Science* 291, 5501 (2001), 52–54. https://doi.org/10.1126/science.10.1126/SCIENCE.1056960

[6] I. Karth and A. M. Smith. 2017. WaveFunctionCollapse is constraint solving in the wild. *Proceedings of the 12th International Conference on the Foundations of Digital Games* (2017). https://doi.org/10.1145/3102071.3110566

[7] Róisín Loughran and Michael O'Neill. 2020. Evolutionary music: applying evolutionary computation to the art of creating music. *Genetic Programming and Evolvable Machines* 21, 1-2 (feb 2020), 55–85. https://doi.org/10.1007/s10710-020-09380-7

[8] Maxim Gumin. 2016. *WaveFunctionCollapse*. https://github.com/mxgmn/WaveFunctionCollapse

[9] Edward Pearsall. 1997. Interpreting Music Durationally: A Set-Theory Approach to Rhythm. *Perspectives of New Music* 35 (1997), 205–230. https://doi.org/10.2307/833685

[10] Bennett Reimer. 1995. The Experience of Profundity in Music. *Journal of Aesthetic Education* 29, 4 (1995), 1–21. https://doi.org/10.2307/3333288

[11] Martin A. Rohrmeier. 2020. Towards a Formalization of Musical Rhythm. In *International Society for Music Information Retrieval Conference*. https://doi.org/10.5281/ZENODO.4245508

[12] Ryohey. 2023. Signal. https://signal.vercel.app/

[13] Matthew D. Schulkind. 1999. Long-term memory for temporal structure:. *Memory Cognition* 27, 5 (sep. 1999), 896–906. https://doi.org/10.3758/bf03198542

[14] Noor Shaker, Julian Togelius, and Mark J. Nelson. 2016. *Procedural Content Generation in Games*. Springer International Publishing. https://doi.org/10.1007/978-3-319-42716-4

[15] Pál Patrik Varga and Rafael Bidarra. 2023. ProceduraLiszt. https://bit.ly/proceduraliszt_app

[16] Pál Patrik Varga and Rafael Bidarra. 2024. Harmony in Hierarchy: Mixed-Initiative Music Composition Inspired by WFC. Submitted for publication.

[17] Haili You. 1994. Defining rhythm: Aspects of an anthropology of rhythm. *Culture, Medicine and Psychiatry* 18, 3 (sep 1994), 361–384. https://doi.org/10.1007/bf01379231