Javanese Gamelan Composition Using Modified N-Gram Language Model

Muhammad Luqman Hakim (13523044)
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13523044@std.stei.itb.ac.id muhluqhakim@gmail.com

***Abstract*—This paper introduces a novel variation of the n-gram model tailored to generate Javanese gamelan music, where traditional musical structures are enhanced with positional information to capture the cyclic and hierarchical nature of gamelan compositions. The position of each note within the musical cycle is proposed as a condition for generating subsequent notes, enhancing the accuracy and relevance of the generated sequences. The model is trained on musical datasets that include both the notes and their respective positions within the structure, allowing for a more context-aware prediction of note sequences. The mathematical framework for the modified n-gram model is provided for both training and generation. By combining probabilistic methods with a deeper understanding of the unique properties of gamelan music, this approach aims to improve the quality of computer-generated gamelan compositions while preserving the traditional idiomatic features of the music. The results demonstrate the potential of this model for use in computational musicology and automated composition systems for gamelan music.**

***Index Terms*—Gamelan, n-gram language model**

I. Introduction

Javanese gamelan music, a rich and intricate musical tradition originating in Java, Indonesia, is characterized by its complex rhythmic structures, layered melodies, and cyclical forms. The *balungan*, or the core melody, serves as the foundation of most gamelan compositions, providing a framework that guides the interplay of various instruments. Despite its relatively fixed melodic structure, gamelan music is dynamic, with its patterns evolving over time and influenced by the interactions between musicians. These unique features present challenges for computational music generation, as traditional music models often fail to capture the nuances of such a structured yet fluid tradition.

Probabilistic models, such as the n-gram model, have been widely used in language modeling and music generation, leveraging the statistical dependencies between consecutive elements in a sequence. In music, n-gram models predict the next note based on a fixed-length history of preceding notes, treating them as independent events. However, while n-gram models are effective in modeling local note dependencies, they often overlook important contextual factors such as the position of a note within the larger structure of a composition. In Javanese gamelan music, the position of a note within the cycle plays an important role in determining its pitch, rhythm, and relation to other notes. The cyclical nature of gamelan compositions further complicates the task, as the music often returns to earlier motifs and phrases, influencing the patterns of note sequences.

To address these limitations, I propose a modified version of the n-gram model that incorporates positional information into the generation process. By conditioning the generation of each note not only on the preceding notes but also on its position within the composition, our model more effectively captures the structural dependencies that define the rhythmic and melodic flow of gamelan music. This approach allows for the generation of more contextually relevant musical sequences that respect both local dependencies and global structural patterns.

In this paper, I outline the mathematical framework for the modified n-gram model, and discuss the benefits of incorporating positional information. The purpose of this study is to contribute to the field of computational musicology and to offer a tool for generating authentic, context-sensitive gamelan music that can be used for research, composition and performance.

II. Javanese Gamelan

Javanese gamelan music is a traditional ensemble music form from Indonesia, characterized by its rich textures, cyclical structures, and interlocking patterns. It features a variety of instruments, including metallophones, gongs, drums, and stringed instruments, each playing a specific role within the ensemble. The music is built around a core melody known as the *balungan*, which serves as the foundation for other instruments to elaborate upon. The *balungan* is typically played by metallophones and represents the skeletal structure of the composition, guiding the interplay between rhythmic and melodic layers. There are three types of balungan: *balungan mlaku*, *balungan nibani*, and *balungan rangkep* [1].

Balungan nibani is characterized by its rhythmic structure, where every other beat is left empty, resulting in a density that is half of the regular balungan pulse.

Balungan mlaku, sometimes called *balungan mlampah*, maintains a rhythmic density equal to the regular balungan pulse.

Balungan rangkep, also referred to as *balungan ngadhal* or *balungan tikel*, is a denser form of balungan with a rhythmic pulse that is double the regular balungan pulse.

III. Probability

Probability provides a mathematical framework for quantifying the likelihood of events. Random experiments, such as flipping a coin or rolling a die, produce outcomes that are uncertain but belong to a well-defined sample space S. Each subset of S is called an event, and the probability of an event A is a number between 0 and 1, denoted as $P(A)$. The value 0 represents an impossible event, while 1 indicates certainty.

The classical definition of probability assumes equally likely outcomes and is calculated as:

$$P(A) = \frac{|A|}{|S|},$$

where $|A|$ is the number of favorable outcomes in A, and $|S|$ is the total number of outcomes in the sample space S.

IV. Conditional Probability

Conditional probability measures the likelihood of an event occurring given that another event has already occurred [4]. It provides a way to refine predictions by incorporating known information. Mathematically, the conditional probability of an event A given B is defined as [5]:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

Here, $P(A \cap B)$ represents the probability that both A and B occur simultaneously, while $P(B)$ is the probability that event B occurs. The condition $P(B) > 0$ is necessary to ensure that the conditional probability is well-defined, as it would not make sense to condition on an event with zero probability.

The formula indicates that to compute the conditional probability, one must first determine how often both events A and B happen together and then normalize this by the probability of event B alone. This normalization adjusts the probability to account for the known occurrence of B, effectively narrowing the focus to the cases where B is true. In other words, conditional probability quantifies how the likelihood of A changes when we know that B has occurred.

This framework is particularly powerful in domains where the probability of an event depends on additional contextual information. In natural language processing, for instance, the probability of a word in a sequence can depend on the preceding words, allowing for the prediction of future words based on context. Similarly, in music, the occurrence of certain musical elements, such as the pitch or rhythm of a note, can depend on the preceding notes or the position of the note within a particular section of the composition.

V. N-Gram Language Model

Conditional probability plays a foundational role in word n-gram language models, which are widely used in natural language processing (NLP) to predict the likelihood of word sequences. An n-gram is a contiguous sequence of n words, and these models aim to estimate the probability of a word given its preceding context, typically limited to the previous $n-1$ words [3]. This predictive ability is rooted in the principles of conditional probability.

In the context of text, let $W_1, W_2, \ldots, W_n$ represent a sequence of words. The probability of the entire sequence can be expressed using the chain rule of probability:

$$P(W_1, W_2, \ldots, W_n) = \prod_{i=1}^{n} P(W_i \mid W_1, W_2, \ldots, W_{i-1})$$

This formula illustrates that the probability of a sequence is decomposed into the product of conditional probabilities of each word, given its preceding words.

However, modeling the full history of preceding words $(W_1, W_2, \ldots, W_{n-1})$ is computationally expensive and often infeasible due to the vast number of possible word combinations. To address this, n-gram models make the Markov assumption, which simplifies the dependency by considering only the most recent $n-1$ words. This reduces the conditional probability to:

$$P(W_n \mid W_1, W_2, \ldots, W_{n-1}) \approx P(W_n \mid W_{n-(N-1)}, \ldots, W_{n-1}).$$

For example: - In a bigram model ($N = 2$), the probability of a word depends only on the previous word:

$$P(W_n \mid W_1, W_2, \ldots, W_{n-1}) \approx P(W_n \mid W_{n-1}).$$

- In a trigram model ($N = 3$), the probability of a word depends on the two preceding words:

$$P(W_n \mid W_1, W_2, \ldots, W_{n-1}) \approx P(W_n \mid W_{n-2}, W_{n-1}).$$

The probability of a sequence of words in an n-gram model can thus be computed as:

$$P(W_1, W_2, \ldots, W_n) = \prod_{i=1}^{n} P(W_i \mid W_{i-(n-1)}, \ldots, W_{i-1}),$$

where for the first $n-1$ words, appropriate smoothing or assumptions handle missing context.

Although n-gram models are conceptually simple, their reliance on conditional probability is a powerful example of how probabilistic methods can model and predict linguistic patterns effectively.

VI. Modification to the N-Gram Language Model

In musical compositions, the position of a note within a song carries significant contextual importance. Traditional n-gram models focus solely on the sequential relationship between neighboring notes, capturing dependencies within local windows of $n-1$ preceding elements. However, music often follows structured patterns that evolve over time and across different sections of a composition. The inclusion of positional information adds a temporal dimension to the model, enabling it to account for these broader structural and rhythmic contexts.

In many musical pieces, the placement of notes within phrases, measures, or beats influences their likelihood and musical function. For instance, notes that occur at the beginning of a phrase often serve as introductory elements, while notes at the end are more likely to resolve to stable tones, such as the tonic or dominant. Similarly, strong beats within a measure frequently carry chordal tones, whereas weaker beats may introduce passing or decorative notes. These patterns are not only sequential but also position-dependent, reflecting the hierarchical and temporal nature of musical organization.

By conditioning the probability of a note on its position in the song in addition to the preceding $n-1$ notes, the model can capture both local dependencies and global structural trends. The modified conditional probability can be expressed as $P(w_t \mid w_{t-(n-1)}, \ldots, w_{t-1}, \text{Position}_t)$, where Position_t represents the location of the note in the song. This extension acknowledges that a note's role and likelihood are not solely dictated by immediate neighbors but also by its broader context within the composition. Such an approach is especially valuable for modeling genres with strong rhythmic or thematic patterns, as it allows the model to learn and predict musical structures more effectively. Incorporating position as a condition thus enhances the expressive power of the n-gram model, aligning it more closely with the nuanced nature of musical notation.

VII. Implementation

A. Choice of Programming Language and Library

The program is implemented in Python, without the use of external libraries.

B. Data Structure

The program employs a dictionary data structure to record the frequency of each n-gram encountered in the training data. This approach enables efficient storage and retrieval of frequency counts for various combinations of preceding notes, positions, and predicted notes.

In the program, each beat is represented as a tuple, allowing the model to flexibly encode the number of notes within a single beat. If a beat contains two notes, the tuple includes two elements, each representing one note. Similarly, a beat with only one note is represented as a tuple with a single element, while a beat with four notes is represented as a tuple with four elements.

C. Training Process

The training process constructs a probabilistic model that incorporates both the sequence of notes and their positional information. The input consists of balungan sequences and their corresponding positional annotations, derived from authentic Javanese gamelan compositions. For each sequence in the dataset, the algorithm extracts n-grams. Each n-gram comprises a sequence of $n-1$ preceding notes, the current note to be predicted, and its position within the musical cycle.

The algorithm records the counts of joint occurrences for every combination of preceding notes, positions, and predicted notes. These counts are used to compute conditional probabilities by normalizing the joint occurrence counts with respect to the total occurrences of the corresponding preceding notes and positions. The result is a probability table that defines $P(\text{note} \mid \text{context, position})$, which serves as the foundation for sequence generation.

D. Generation Process

The generation process utilizes the trained model to produce new balungan sequences. Beginning with an initial state defined by $n-1$ notes and a position, the algorithm predicts the next note by sampling from the conditional probability distribution $P(\text{note} \mid \text{context, position})$ derived during training.

After a note is generated, the state is updated by shifting the sequence window to include the newly generated note and incrementing the position within the musical cycle. This iterative process continues until a sequence of the desired length is created. The positional parameter ensures the generated sequence adheres to the cyclical structure of gamelan compositions, capturing both local dependencies and global positional contexts. The output is a complete balungan sequence that aligns with the traditional structural characteristics of Javanese gamelan music.

Algorithm 1: Training the Modified N-gram Model with Positional Information

Input: List T: a list of musical sequences.
Output: Dictionary $Count$: a dictionary recording the frequency of all $n-1$ and n-grams.

for *sequence* $\in$ *training data* **do**
  for *note* N_t *at position* p_t **do**
    $\text{Count}\big(N_{t-(n-1)}, \ldots, N_t, \text{Position}_t\big)$ += 1
    $\text{Count}\big(N_{t-(n-1)}, \ldots, N_{t-1}, \text{Position}_t\big)$ += 1
  end
end

Algorithm 2: Generating New Music Notation Using the Modified N-gram Model

Input: Dictionary D: a dictionary recording the frequency of all $n-1$ and n-grams.
L: length of musical sequence to be generated
Output: Musical sequence S.

1. S ← random starting sequence
2. **for** *each step in the generation* **do**
  N_t ← random note with probability

$$P(N_t \mid N_{t-(n-1)}, \ldots, N_{t-1}, \text{Position}_t)$$

  S += N_t.
end

VIII. Dataset

The dataset used for this study is sourced from gamelanbvg.com, a resource dedicated to providing comprehensive collections of Javanese gamelan compositions and notations. This dataset serves as the foundation for training the modified n-gram model, offering a rich corpus of traditional balungan notations and their corresponding structures. By leveraging this dataset, the program ensures that the generated balungan sequences align closely with the stylistic and structural characteristics of authentic Javanese gamelan music.

but for the purpose of this study, the dataset is limited to composition in the *Ladrang* form, and with *laras pelog pathet lima*. This limitation is imposed to ensure the accuracy of the model, as different musical forms within gamelan may have distinct structural characteristics. The Ladrang form, with its specific colotomic structure and rhythmic subdivisions, provides a consistent and well-defined dataset for training the model. Using only laras pelog pathet lima further narrows the scope to a particular tuning system, ensuring that the generated balungan sequences align with the traditional tonal and rhythmic conventions of Ladrang compositions in this tuning. Although the program can be applied to other forms, extending the dataset to include them at once introduce variability in structure, which could impact the accuracy and reliability of the generated sequences.

In this study, the dataset is divided into two parts, the *buka* and the section that follows it, for the purpose of training the model more effectively. The reason for this division is that the *buka* serves as an introductory section with a distinct musical structure and simpler patterns compared to the main body of the composition. By training the model on the *buka* separately, it allows the model to learn the specific characteristics of the introductory material, such as its rhythmic and melodic patterns. Afterward, the model is trained on the remaining section, which follows a more structured form and involves a fuller, more complex arrangement of instruments and rhythmic patterns. This separation ensures that the model can capture the unique features of both the *buka* and the main section, improving its overall ability to generate or predict the musical patterns in each part more accurately. Hence, training the model on the *buka* and the following section separately helps to avoid confusion between the simpler introductory material and the more intricate patterns of the main composition.

IX. Result

$$.\,\underset{\cdot}{6}\underset{\cdot}{5}\underset{\cdot}{4}\quad \underset{\cdot}{2}\underset{\cdot}{4}\underset{\cdot}{6}\underset{\cdot}{5}\quad 77..\quad 1232$$

$$5654\quad 2123\quad 1132\quad .\,165$$

$$.\,654\quad 2121\quad .111\quad 2353$$

$$.\,523\quad 5123\quad 5321\quad \underset{\cdot}{6}\underset{\cdot}{5}\underset{\cdot}{3}\underset{\cdot}{5}$$

The notation above represents an example of generated music notation derived from the programmatic application of the modified n-gram model. Each number corresponds to a note or set of notes for a single beat in the musical structure of Javanese gamelan, specifically within the *Ladrang* form, in *laras pelog pathet lima*.

In this context, the generated notation captures the skeletal melody (*balungan*), which serves as the foundation for elaboration by other instruments in the ensemble.

The musical sequence represents the output of the program, which combines statistical modeling and the syntactical rules of gamelan music to generate plausible melodic patterns. This approach ensures that the notation adheres to the cultural and structural conventions of the Javanese *Ladrang* form while providing a programmatic framework for automatic music generation.

X. Conclusion

The application of a modified n-gram model to generate *balungan* notation for Javanese gamelan music demonstrates the potential of computational methods in modeling and synthesizing traditional music. By incorporating both sequential dependencies and positional information, the model effectively captures the nuanced structures characteristic of the *Ladrang* form in the *laras pelog pathet lima* tuning system. The use of statistical modeling allows the generated music to adhere closely to the idiomatic patterns observed in traditional gamelan compositions, while the inclusion of detailed features, such as positional data and beat-wise note tuples, ensures fidelity to the cyclical and hierarchical nature of the music.

The dataset's focus on a specific musical form and tuning system provides a controlled environment for accurate modeling, though the program's design retains the flexibility to adapt to other forms and tunings if appropriately trained. This adaptability highlights the program's broader applicability, offering opportunities for exploring various gamelan styles or even other forms of traditional ensemble music. The generated notation preserves essential elements of Javanese gamelan's cultural and musical identity, making it a valuable tool for both academic study and creative exploration.

Future work could explore integrating additional features, such as dynamics, ornamentation, or variations in *irama*, to further enhance the authenticity and richness of the generated compositions. By bridging computational techniques with the intricacies of traditional music, this approach contributes to the ongoing dialogue between technology and cultural heritage, ensuring the preservation and evolution of Javanese gamelan music in modern contexts.

XI. Acknowledgment

I would like to express my sincere gratitude to God Almighty for His guidance, who made the completion of this paper possible. I would also like to thank Dr. Ir. Rinaldi Munir, M.T., for his invaluable guidance and support during the IF2123 Aljabar Linier dan Geometri course. Additionally, the completion of this paper has been greatly supported by my dear family and friends, fellow students from the Teknik Informatika and Sistem dan Teknologi Informasi departments at ITB, Class of 2023.

References

[1] Drummond, B. (n.d.). JAVANESE GAMELAN TERMINOLOGY. gamelanbvg.com. Retrieved January 8, 2025, from https://gamelanbvg.com/gendhing/gamelanGlossary.pdf

[2] Sorrell, N. (1990). A guide to the gamelan. https://ci.nii.ac.jp/ncid/BA11049005

[3] Bengio, Y., Ducharme, R., Vincent, P., & Janvin, P. (2003). A neural probabilistic language model. Journal of Machine Learning Research. https://doi.org/10.5555/944919.944966

[4] Gut, A. (2005). Probability: a graduate course. In Springer texts in statistics. https://doi.org/10.1007/b138932Gut, A. (2005). Probability: a graduate course. In Springer texts in statistics. https://doi.org/10.1007/b138932

[5] Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., Meester, L. E. (2005). A modern introduction to probability and statistics. In Springer texts in statistics. https://doi.org/10.1007/1-84628-168-7

Appendix A

Complete Implementation of the Compression Program

The complete implementation of the program can be accessed at https://github.com/carasiae/gamelan-skibidi

STATEMENT

I hereby declare that the paper I wrote is my own writing, not an adaptation, or translation of someone else's paper, and not plagiarized.

Bandung, 2 January 2025

Luqman

Muhammad Luqman Hakim