Benzaiten: A Non-expert-friendly Event of Automatic Melody Generation Contest

Yoshitaka Tomiyama[1,2], Tetsuro Kitahara[3], Taro Masuda[1,4], Koki Kitaya[1],
Yuya Matsumura[1,5], Ayari Takezawa[1,6], Tsuyoshi Odaira[7], and Kanako Baba[8⋆]

[1]Music×Analytics Meetup Steering Committee,[2]NABLAS Inc.,[3]Nihon University,[4]Nikkei Inc.
[5]NOVASELL Inc.,[6]Kyushu University,[7]Freelance,[8]Advanced Institute of Industrial Technology
musicanalyticsmeetup@gmail.com

Abstract. This paper presents a contest-style music evaluation event called *Benzaiten*. There have been some attempts to evaluate different music generation systems with a unified criterion and/or platform, but it was not an event that non-experts could easily enjoy. At Benzaiten, we encouraged non-researcher people to join it as entrants by providing starter kits and communication channels. As well, we exercised ideas towards a high-quality entertainment event for laypeople to enjoy it. As a result, 15 people joined this event as entrants (eight of which moved to the main round), and more than 100 people participated as the audience.

Keywords: Melody generation, evaluation, contest

1 Introduction

Whereas the research on automatic melody generation has a long history, the recent development of machine learning (ML) technologies has been rapidly increasing the number of attempts at automatic melody generation [1].

Evaluating melody generation systems/methods is still an important open problem in this field. Unlike speech recognition and image classification, the *correct* output (e.g., a melody) for a certain input (e.g., a chord progression) cannot be uniquely or objectively determined. We, therefore, have to conduct subjective quality tests on generated melodies employing music experts, but its methodology has not been necessarily established.

To provide a platform for evaluating different systems/methods on a unified criterion, some researchers made attempts to organize contest-based evaluation. Sturm et al. [2] organized the AI Music Generation Challenge 2000, in which they collected Irish double-jig pieces from entrants and hired experts to evaluate them. They also organized the 2021 edition focusing on Swedish traditional dance music [3]. Katayose et al. [4] held Performance Rendering Contests (Rencon) to provide a subjective evaluation platform for researchers developing expressive music performance rendering systems. Yeh
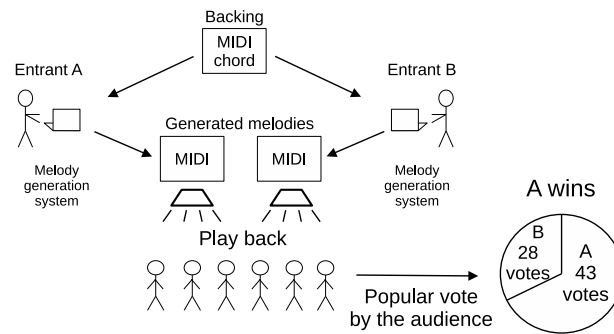

Fig. 1. Basic scheme of Benzaiten. Two entrants are given a backing track and generate melodies during the event. Then, the winner is determined according to popular votes by the audience.

et al. [5] attempted on collaborative comparisons of harmonization systems developed by researchers from different institutes. However, these attempts have the following two problems:

- Because entrants are implicitly assumed to have skills or experiences in developing music generation systems/methods, there is no scheme for encouraging non-researcher people who have not tried such development.
- Because the primary purpose is to provide a unified platform for evaluation, they are not necessarily fun for laypeople as entertainment shows.

In this paper, we propose a novel contest-based music generation evaluation event, called *Benzaiten*. The most important policies in Benzaiten are *openness* and *fun* for non-experts. To develop the melody generation field furthermore, it is important to let a wide range of people have an interest in it. We, therefore, aim at an event that non-experts can enjoy as entrants and/or the audience. Benzaiten has the following features:

- To make it easy for novices to join it as entrants, we provided starter kits for developing melody generation systems and communication channels on Slack for sharing issues and ideas among potential entrants.
- To make it possible for laypeople to enjoy it as the audience, we exercised some ideas to make its quality as an entertainment show higher, including a popular vote and a one-on-one battle scheme.

2 Basic policy and event design

Benzaiten (Figure 1) is a contest-based melody generation evaluation event. Every entrant brings their melody generation system and generates a melody that fits a given backing track. The generated melody is played back within the event, then the winner is determined based on voting. As discussed in the Introduction, the basic policies in this event are *openness* and *fun* for non-experts as follows:

- *Openness to novices*: It is easy for various people to join this event as entrants, even if they are not ML and/or music experts.
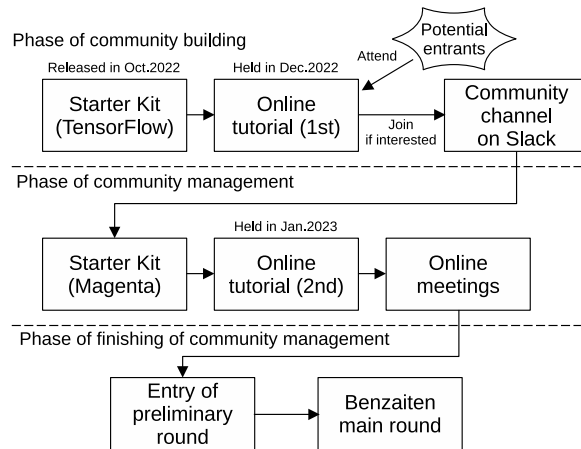
Fig. 2. Benzaiten's community management

– *Fun for laypeople*: Participants can enjoy this event as the audience even if they do not have music- or ML-related knowledge.

To achieve *openness to novices*, we make the following attempts:

– **Starter kits**: We developed two kinds of starter kits for this event, with which everyone can quickly try automatic melody generation.
– **Online tutorials**: We held online tutorials, in which the tutors taught the basic knowledge of MIDI and music as well as how to use the starter kits.
– **Community management of potential entrants**: We made a community of potential entrants, including those who had yet to determined entry. We encouraged communication among them by making a Slack channel and holding online meetings.

To achieve *fun for laypeople*, we introduce the following ideas to the event:

– **Popular vote**: The audience can get involved in determining the winner.
– **One-on-one battle scheme**: We adopted a tournament style based on a one-on-one battle scheme. This scheme makes the voting for each match simple, because all the participants have to do is to judge which is better of the presented two melodies.
– **Live melody generation**: For each match, the backing track is provided right when the match starts. Therefore, the entrants must generate melodies live during the event (they cannot generate melodies in advance).

3 Actions before the event

To encourage a wide range of people to join the event as entrants, we should promote this event widely and encourage as many people as possible to join the community of potential entrants. Therefore, we did the following (Figure 2).
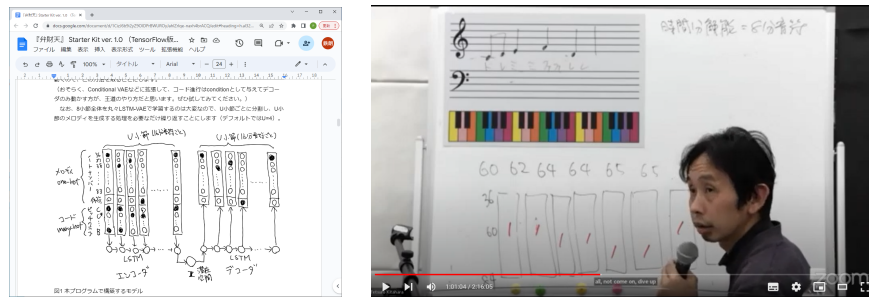
Fig. 3. Starter Kit (TensorFlow Edition) (left) and its tutorial (right)

3.1 Release of Starter Kit (TensorFlow Edition)

To encourage people who are not familiar with automatic melody generation, we developed the Starter Kit TensorFlow Edition (Figure 3.1 Left)[1], in which users can easily try melody generation based on LSTM-VAE. This starter kit provides the codes for:

– Reading MusicXML files taken from Charlie Parker's Omnibook MusicXML Data [6] and extracting melodies from them,
– Converting melodies to sequences of one-hot vectors,
– Training an LSTM-VAE model with prepared melodies using TensorFlow, and
– Generating a melody with the trained LSTM-VAE model in the MIDI format.

All codes are provided under the MIT License. As well, the starter kit provides several hints for extending those codes.

3.2 Tutorial on Starter Kit (TensorFlow Edition)

We conducted a tutorial aiming at allowing potential entrants to quickly understand the Starter Kit (TensorFlow Edition) on Zoom (Figure 3.1 Right). We encouraged a wide range of people to participate in this tutorial, including those yet to determined to join Benzaiten as entrants. 124 people participated, and some of them expressed their interest in the entry. This tutorial was well-received because we explained general knowledge of MIDI, the harmony theory as well as the codes of the starter kit.

3.3 Encouraging discussions on Slack

Because several participants of the above-mentioned tutorial expressed interest in joining Benzaiten as entrants, we invited them to the discussion channel for the community of potential entrants on Slack. We intended to let them exchange various information with each other and to foster a mood of friendly rivalry. They actually discussed some topics including: how to find publicly available datasets, and how to execute the starter kit's codes on a local computing resource with MacOS on an M1 chip.

---

[1] Available at the following URL (written in Japanese): `https://docs.google.com/document/d/1CizJ6b9i2yZ9OIDPrBWUROyJahlZrlqe-naxh4brACQ/`

3.4 Starter Kit (Magenta Edition) and its tutorial

We released the Starter Kit (Magenta Edition)[2], which illustrates how to generate melodies using ImprovRNN, included in Magenta [7]. Because this edition uses the pre-trained ImprovRNN model, unlike the TensorFlow edition, users can try melody generation more simply, even though improving the model is complex. Furthermore, we held a tutorial for explaining this starter kit. 13 potential entrants participated in this tutorial.

3.5 Online meetings for progress sharing

We held online meetings twice to enable potential entrants to share each other's progress. 16 people participated in the first meeting, and 13 participated in the second one.

4 Preliminary round

15 people (or teams) entered Benzaiten even though the number of acceptable entrants was limited to eight. We, therefore, held an online preliminary round. On the designated web page, participants listened to all melodies submitted by 15 entrants and rated them on a scale of 0 to 10. To avoid bias caused by the order etc., the web page lists the melodies anonymously in a random order. We promoted this preliminary round on Twitter, and ratings by 70 participants were collected. Finally, the eight entrants with high ratings moved into the main round.

5 Implementation and results of Benzaiten (main round)

5.1 Outline of the event

Benzaiten adopts a single-elimination tournament style. Because we accepted eight entrants, the event consists of seven matches: four quarterfinals, two semifinals, and one final. The overall schedule is as follows:

14:00–14:10 Opening
14:10–15:30 Four quarterfinals (20 mins for each match)
15:30–16:00 Lightning talks by all entrants on the techniques they use
16:00–16:40 Two semifinals
16:40–16:55 Sponsors' lightning talks
16:55–17:15 The final
17:15–17:35 Long talk by T. Kitahara
17:35–18:00 Awards ceremony & Ending

Each match was conducted as follows (Figure 7):

1. The two entrants go on the stage.
2. The organizer plays back the backing track for that match.

---

[2] Available at the following URL (written in Japanese): `https://colab.research.google.com/drive/1isnq_E2Mc-Fzeb8DKzYGL39l-B30AwZK`

Fig. 4. Scene during a match (left) and winner announcement (right)

3. The backing track data (a MIDI file, a chord transcription file) is put into the online storage. The entrants are not permitted to download them before the match starts.
4. The match starts. The entrants must submit the generated melody's MIDI file before five minutes pass.
5. The melodies submitted by the two entrants are played back.
6. The popular vote starts. Google Forms is used as a voting platform.
7. The voting result (the winner) is announced.

5.2 Contest rule

To achieve a fair contest, we carefully designed the contest rule and presented the rule book on the Web. A distinctive rule is to allow entrants to compare more than one generated melody by listening and choose one within a five-minute time limit. We also decided the length and the timbre (program change) of melodies; The entrants cannot change them to focus on the quality of the melodies. Chord transcriptions are given as a text file.

5.3 Backing tracks

To make this event successful, we consider it important to present high-quality backing tracks that include a variety of chord progressions. To achieve this, a professional musician joined our team and composed seven high-quality backing tracks with different chord progressions.

According to the contest rules, every backing track has nine measures with the key of C major or A minor. The used chord progressions and keys are listed in Table 1. Backing tracks for Matches 1, 4, and 6 contain chords with non-diatonic roots, which make appropriate melody generation slightly difficult.

The chord transcription file (in the CSV format) as well as the MIDI file of the backing track are given to the entrants at each match.

5.4 Entrants

The eight entrants are listed in Table 2. Two entrants (*log5* and *Dekoboko Friends*) improved the post-processing to avoid musically unnatural notes without improving the starter kits' ML model. On the other hands, two entrants (*yatszhash* and *nayopu*) adopted completely different approaches, that is, melody generation by ABC-notated text generation or notation image generation. The other entrants adopted well-known ML models such as a Conditional VAE, a CNN-VAE, and a Transformer.

Table 1. Chord progressions and keys of backing tracks

Match	Round	Chord progression								Key
1		`C`	`|G/B`	`|Bb`	`|F/A`	`|Fm/Ab`	`|C/G`	`|D/F#`	`|G` `|C`	C maj
2	Quarter-	`C`	`|C/E`	`|F`	`|G E`	`|Am`	`|C/G`	`|D7/F#`	`|Dm7G7` `|CM7`	C maj
3	finals	`Am`	`|F`	`|G`	`|C G`	`|Am`	`|F`	`|G`	`|E7` `|Am`	A min
4		`C`	`|G`	`|G7`	`|C`	`|C`	`|Bb`	`|F`	`|C` `|C`	C maj
5	Semi-	`Am`	`|DmEm`	`|Am7`	`|DmEm`	`|Am7`	`|DmEm`	`|FM7 G`	`|Am` `|Am`	A min
6	finals	`Am`	`|Am/F#`	`|Dm`	`|Bb`	`|Am`	`|Am/F#`	`|Dm`	`|Dm/B E7` `|Am`	A min
7	Final	`Dm7 F/G`	`|CM7 FM7`	`|Dm/B E7`	`|A7`	`|`				A min
		`Dm7 F/G`	`|CM7 FM7`	`|Dm/B E7`	`|Am`	`|Am`				

Table 2. Eight entrants and their melody generation techniques

No.	Name	Used techniques	Result
1	yatszhash	Generated a ABC-notated text with a language model	
2	log5	Original post-processing based on music theory (with the un-modified ML model of the Starter Kit (TensorFlow))	Champion
3	T. N.	Modified the Starter Kit (TensorFlow) to CNN-VAE	
4	nayopu	Generated notation images with DALL-E and converted it to the MIDI format	Semifinalist
5	AJI	Implemented MusicTransformer with their original dataset	
6	M. Y.	Generated melodies with ImprovRNN and then reconstructed them with MusicVAE	Semifinalist
7	konumaru	Modified the Starter Kit (TensorFlow) to Conditional VAEs.	
8	Dekoboko Friends	Implemented post-processing to add grace notes and glissando (with the unmodified Starter Kit (Magenta))	Runner-up

5.5 Number of participants

133 people (including the organizers) participated in Benzaiten on Zoom and about 40 to 50 people (including the organizers and entrants) participated on site. At each match, 54.7 votes were collected on average (max: 61, min: 51).

6 Discussions

6.1 Did we succeed in encouraging non-experts to participate?

Out of the eight entrants, five were non-music-related data scientists; they dealt with automatic MIDI generation for the first time. This fact shows that Benzaiten was able to reach out to a wide range of non-experts.

6.2 Did community management work well?

Our community management allowed potential entrants to exchange various information including: codes for porting the starter kit to a local environment, publicly available MIDI datasets, and chord notation.

Sharing MIDI data generated by entrants at online meetings enabled us to check if the MIDI data met our rules. It was effective to avoid errors occurring when we played them back during the event. In fact, no entrants generated erroneous MIDI data.

6.3 Technical trends in entrants

As we mentioned, entrants had various approaches ranging from improvements of the starter kits' post-processing (*log5*, *Dekoboko Friends*) to completely novel ones (*nayopu*, *yatszhash*). The voting results showed that relatively *safe* melodies tended to win. In fact, melodies with some dissonant notes tended to be evaluated low.

Melodies generated by *log5* (the winner) and *Dekoboko Friends* (the runner-up) had different tendencies. The former was a sequence of close-packed short (e.g., 16th) notes, sometimes like arpeggios. The latter consisted of multiple phrases including rests between the phrases. Their system generated multiple melodies, and they selected one that included both notes and rests in a balanced way by checking them via piano rolls.

6.4 Future challenge

One issue in our community management was that discussions among the entrants had not gathered momentum until the date of Benzaiten was approaching. On the other hand, at Kaggle [8], an online platform for data science competitions, entrants can identify their current ranks through the leaderboard. It gives them motivation for continuous improvement. Also in Benzaiten, we need a mechanism like the leaderboard, which makes entrants' current ranks public. If Benzaiten has such a mechanism, entrants should have more motivation for continuously improving their melody generation models, and hence they would participate in the discussion more actively.

7 Conclusion

In this paper, we presented an automatic melody generation contest called *Benzaiten*. This event's features are to encourage non-experts to join it as entrants with starter kits and community management and to aim at a high-quality entertainment event to allow laypeople to enjoy it. Through these attempts, we let many non-expert people have interest in automatic melody generation. In the future, we would like to extend this attempt to let various people, ranging from hobbyists and musicians to ML researchers, begin music generation development. We believe that it would bring remarkable findings, contributing to further progress in our music generation field.

References

1. J.-P. Briot, G. Hadjeres, and F.-D. Pachet: Deep Learning Techqniuques for Music Generation — A Survey, arXiv:1709:01620, 2017.
2. B. L. T. Sturm and H. Maruri-Aguilar: The Ai Music Generation Challenge 2020: Double Jigs in the Style of *O'Neill's 1011*, *J. Creat. Music Syst..* 5(1), 2021.
3. B. L. T. Sturm: The Ai Music Generation Challenge 2021: Summary and Results, *Proc. AIMC 2022*, 2022.
4. H. Katayose et al.: On Evaluating Systems for Generating Expressive Music Performance: the Rencon Experience, *J. New Music Res.*, 41(4), pp.299–310, 2012.
5. Y.-C. Yeh et al.: Automatic Melody Harmonization with Triad Chords: A Comparative Study, *J. New Music Res.*, 50(1), pp.37–51, 2021.
6. https://homepages.loria.fr/evincent/omnibook/
7. https://magenta.tensorflow.org/
8. https://www.kaggle.com/