An Interpretable, Flexible, and Interactive Probabilistic Framework for Melody Generation: Supplementary materials

1 Supplementary Materials Contents

In our supplementary materials, we provide our full survey instrument including all musical excerpts, a discussion of our ablation experiments in Section 2 with musical excerpts, and a README file describing the use of SchenkComposer. We plan to release our complete, reproducible code upon acceptance. We also provide additional figures and statistics for our survey experiments in Section 3.

2 Ablation Experiments

We performed ablation experiments on SchenkComposer in which parts of the framework were purposefully sabotaged. Specifically, we produce several samples from SchenkComposer given "boring" or "ugly" harmonic progressions, middleground melodies, middleground harmonic rhythms, phrase structures, or *smoothness* measures. Samples of these experiments are found in the "Ablation Experiment" folder of our supplementary materials. These experiments give qualitative insight into how each component affects the final melody generation.

When sabotaging the harmonic progression, we use progressions that are highly repetitive (I-I-I-I-I) or nonfunctional in common practice (iii-viio7/ii-V-viio). As expected, when the progression is repetitive, the melody also becomes repetitive, with slight variations around the same few notes. Nonfunctional progressions are off-putting, but SchenkComposer is able to create a relatively coherent melody thanks to its Schenkerian underpinnings. It is able to generate somewhat smooth voice-leading.

Providing an unusual middleground melody has perhaps the most interesting effect. Again we experiment with repetition, this time with dissonant and consonant notes. For instance, in C major, SchenkComposer might produce the progression, $I - IV - ii - V - I$; we place a middleground melody of $C - C - C - C - C$ or $D - D - D - D - D$. Because the foreground mostly "snaps" into harmonic tones, the product is still melodically coherent. These sort of pedal tones do occur in tonal music, so it does not necessarily sound too unusual. We also try using completely dissonant middleground notes (notes that do not fit in the given harmony). SchenkComposer can generally work around these dissonances, often making them sound like accented non-harmonic tones. Finally, we enforce a middleground melody with large leaps between notes. SchenkComposer is able to fill in the gaps with moderate success, creating what sounds like a compound melody.

Using a particularly high value for the *smoothness* factor (such as 5) has a jarring effect. The foreground melody randomly jumps to unexpected registers and there is little that SchenkComposer can do to smooth out the melody. This outcome is expected and makes it clear that the smoothness factor is a vital parameter. Thanks to the middleground melody, however, there is still a sensible thread of music for the listener to grab on to.

For the harmonic rhythm, we generated melodies that change harmony every measure, every beat, and at unusual asymmetrical beats. Changing once every measure is quite natural, if not a bit dull after too long. Changing every beat leads to a particularly smooth melody, for it cannot leap between middleground harmonies. This smooth melody also limits the non-harmonic tones available. The model is easily able to cope with asymmetrical beats.

Lastly, we experiment with unusual phrase structures. We try asymmetrical subphrase proportions and ending without a common cadence. The former leads to mixed results depending on the type of asymmetry. We hypothesize that the variance of opinion for asymmetrical subphrase proportions would be very large based on one's musical preferences. Ending without a common cadence is especially jarring, for the melody ends without a feeling of closure. Otherwise, the melodic content is coherent.

3 Additional Survey Analysis

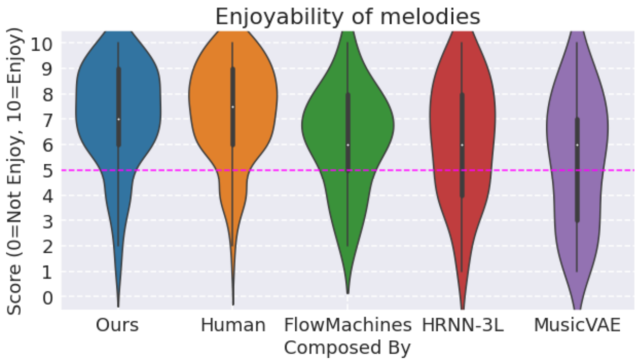

Figure 1: “Enjoyability” of melodies by various models on a scale of 0-10.

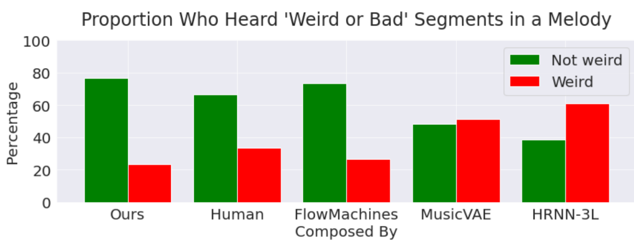

Figure 2: Percentage of participants who heard “weird or bad” segments in melodies by various models.

Question	Low	Medium	High
Years Studied	n=40	27	10
Hours Listened	9	43	25

Question	Nonsense	Wrong	Correct
Meter	13	21	43
Interval	20	38	19

Table 1: Number of participants who answered background and knowledge question. For years studied: low="I have not studied music with a teacher or in school," medium="I have studied music for 5 years or less," high="I have studied music for more than 5 years." For hours listened: low="I listen to music less than 1 hour per week," medium="I listen to music between 1 and 15 hours per week," high="I listen to music more than 15 hours per week."

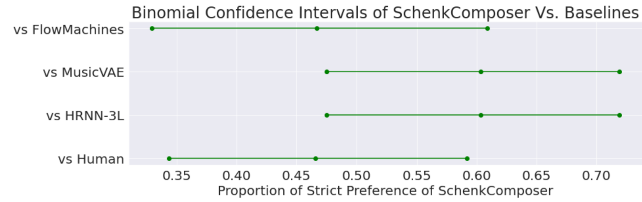

Figure 3: Binomial confidence intervals for strict preference of SchenkComposer over baseline models (higher is better).

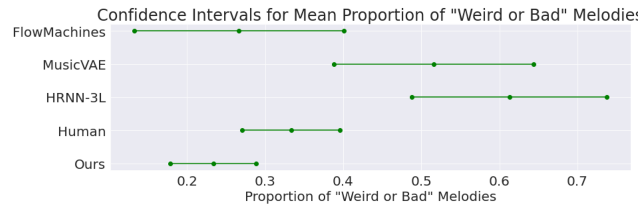

Figure 4: Confidence intervals for “weird or bad” melodies (lower is better).

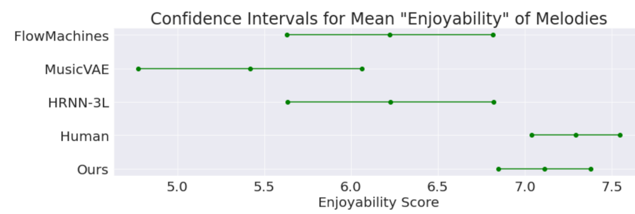

Figure 5: Confidence intervals for enjoyability scores (higher is better).

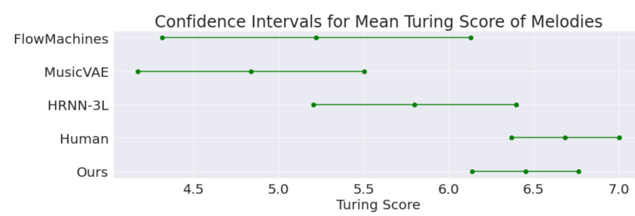

Figure 6: Confidence intervals for Turing scores (higher is better).

4 Genre Experiments

For our genre experiments, we attempt to adjust the parameters of SchenkComposer to generate melodies of specific genres. We attempt to generate style-specific melodies without changes to the contour probabilistic context-free grammar. We are also forced to use a Markov chain for harmony generation due to our lack of hierarchical data in other genres. Despite this lack of hierarchical analyses in other genres, we are able to achieve moderate success. Recordings can be found in the Genre experiments folder.

4.1 Rock/Pop

For the rock/pop genre we simply change the harmonic transition matrix, which may be visualized in Figure 7. Additionally, the foreground rhythms are changed to include more syncopated patterns.

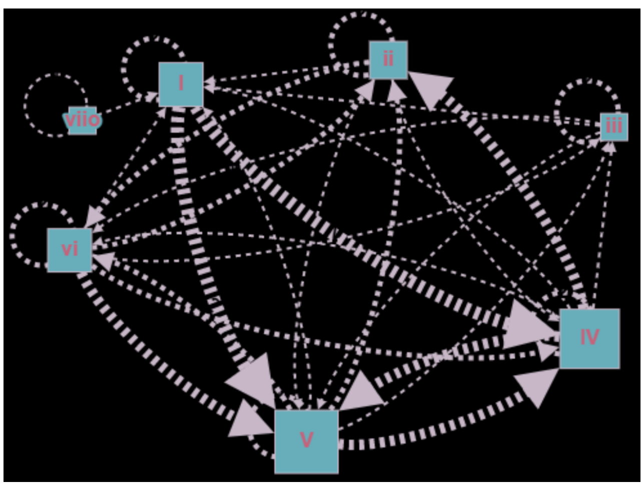

Figure 7: Rock transition graph. Arrows point to harmonies that precede the source harmony. Arrow width indicates the relative probability from one harmony to another.

4.2 Pentatonic

Using the major pentatonic scale, we attempted to generate melodies of the traditional Chinese style. We replace the transition matrix with a simple four-harmony system that only includes notes of the pentatonic scale (see Figure 8).

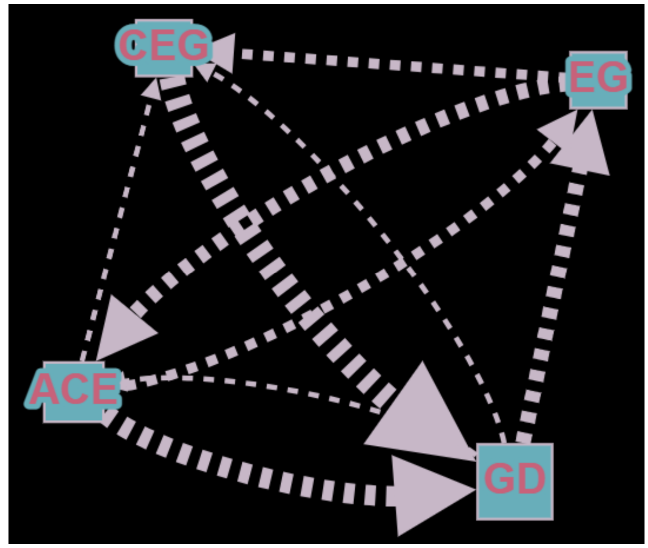

Figure 8: Pentatonic transition graph. Arrows point to harmonies that precede the source harmony. Arrow width indicates the relative probability from one harmony to another.

4.3 Gagaku

Out of sheer curiosity, we attempted to generate melodies in the style of Japanese Gagaku music, although it is far from the authors' expertise. The foreground rhythms were slowed down to last up to multiple measures. We used a four-harmony transition matrix including *Bo*, *Otsu*, *Gyo*, and *Ichi*. Because we are unaware of the transition probabilities, the weights between all harmonies were equal. The melody was based on a subset of the Dorian scale. We based these decision on the information in this website: https://gagaku.stanford.edu/en/.