

MeloForm: Generating Melody with Musical Form based on Expert Systems and Neural Networks

Peiling Lu, Xu Tan, Botao Yu, Tao Qin, Sheng Zhao, Tie-yan Liu

Microsoft Research Asia, Nanjing University, Microsoft Azure Speech

Abstract

Human usually composes music by organizing elements according to the musical form to express music ideas. However, for neural network-based music generation, it is difficult to do so due to the lack of labelled data on musical form. In this paper, we develop MeloForm, a system that generates melody with musical form using expert systems and neural networks. Specifically, 1) we design an expert system to generate a melody by developing musical elements from motifs to phrases then to sections with repetitions and variations according to pre-given musical form; 2) considering the generated melody is lack of musical richness, we design a Transformer based refinement model to improve the melody without changing its musical form. MeloForm enjoys the advantages of precise musical form control by expert systems and musical richness learning via neural models. Both subjective and objective experimental evaluations demonstrate that MeloForm generates melodies with precise musical form control with 97.79% accuracy and outperforms baseline systems in terms of subjective evaluation score by 0.75, 0.50, 0.86 and 0.89 in structure, thematic, richness and overall quality, without any labelled musical form data. Besides, MeloForm can support various kinds of forms, such as verse and chorus form, rondo form, variational form, sonata form, etc.

Musical form Example



Figure 1. The score of the melody part from “Minuet in G Major” by the composer Christian Petzold. This is an example of a melody with musical form as $A(a_1, a_1)B(b_1, b_2)$. We use capital letters (e.g., A, B) to label sections, while using lowercase (e.g., a_1, b_1, b_2) for representing phrases. Different phrases in the same section are labeled with different numbers.

Results

		Structure↑	Thematic↑	Richness↑	Overall↑
Dataset	LMD [40]	3.18 (±0.64)	3.07 (±0.58)	3.14 (±0.40)	3.00 (±0.57)
	POP909 [41]	4.06 (±0.49)	3.83 (±0.54)	4.00 (±0.61)	4.11 (±0.46)
Method	Music Transformer [3]	3.00 (±0.76)	3.21 (±0.74)	2.11 (±0.46)	2.32 (±0.54)
	MELONS [8]	3.18 (±0.64)	3.07 (±0.58)	2.64 (±0.64)	2.89 (±0.52)
	POP909_lm	2.61 (±0.62)	2.93 (±0.60)	2.96 (±0.45)	2.96 (±0.49)
	MeloForm	3.68 (±0.35)	3.57 (±0.36)	3.43 (±0.43)	3.61 (±0.34)

Table 1. Subjective evaluation results, with mean opinion scores and 95% confidence interval for each metric.

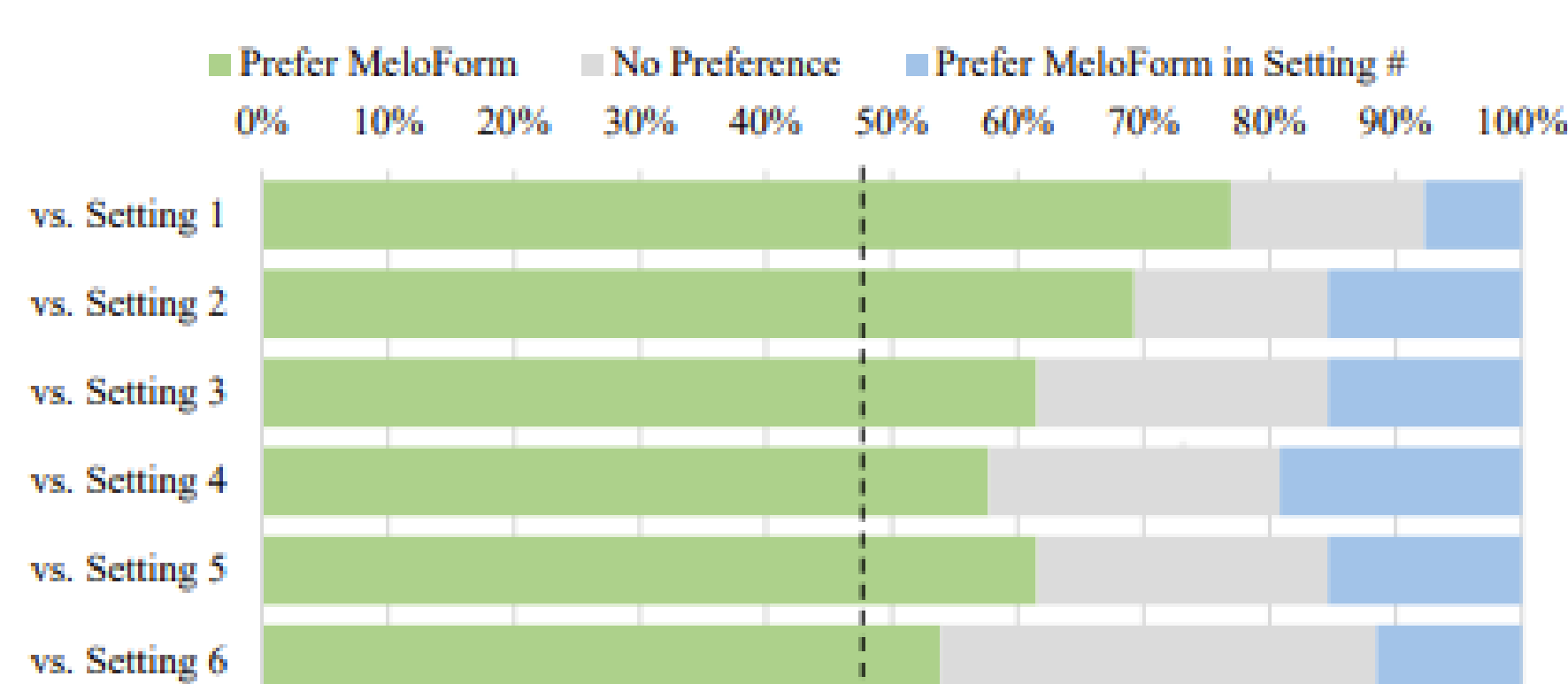
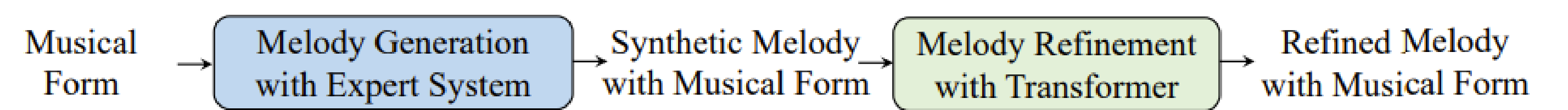


Figure 3. Preference distribution for ablation study, which compares the melodies generated by MeloForm with that from modified system (i.e., MeloForm in Setting #).

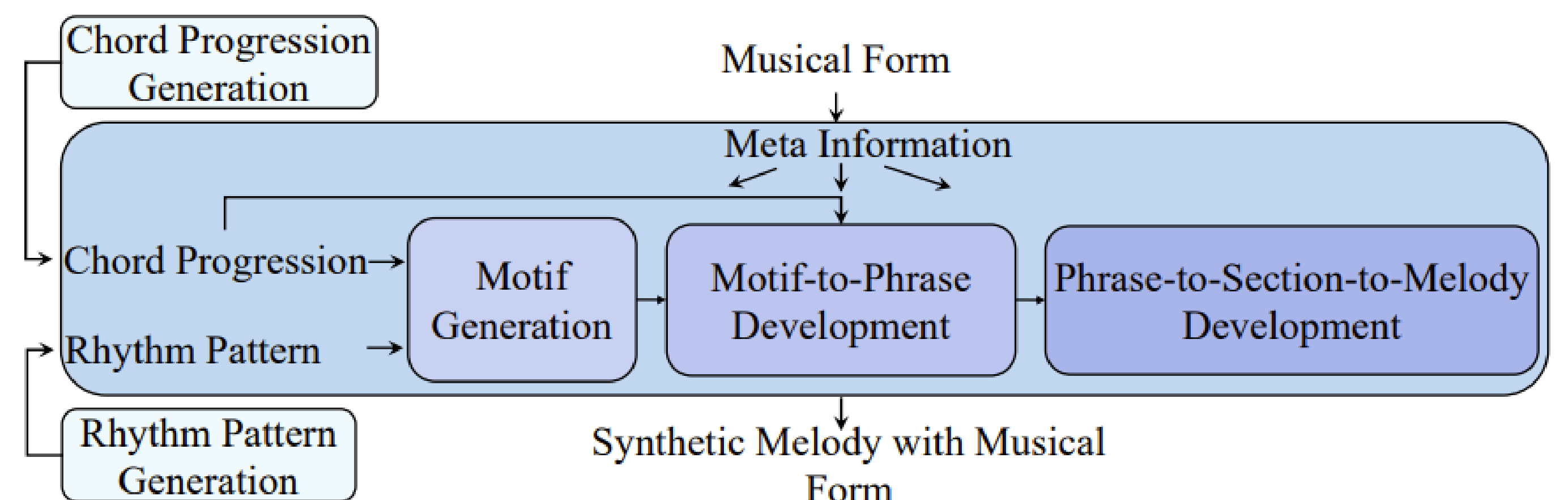
- Setting 1: w/o development strategies from expert systems
- Setting 2: w/o expert systems
- Setting 3: w/o neural networks
- Setting 4: w/o fine-grained rhythm pattern condition
- Setting 5: refinement strategy
- Setting 6: section differentiation

Musical form control accuracy: 97.79%

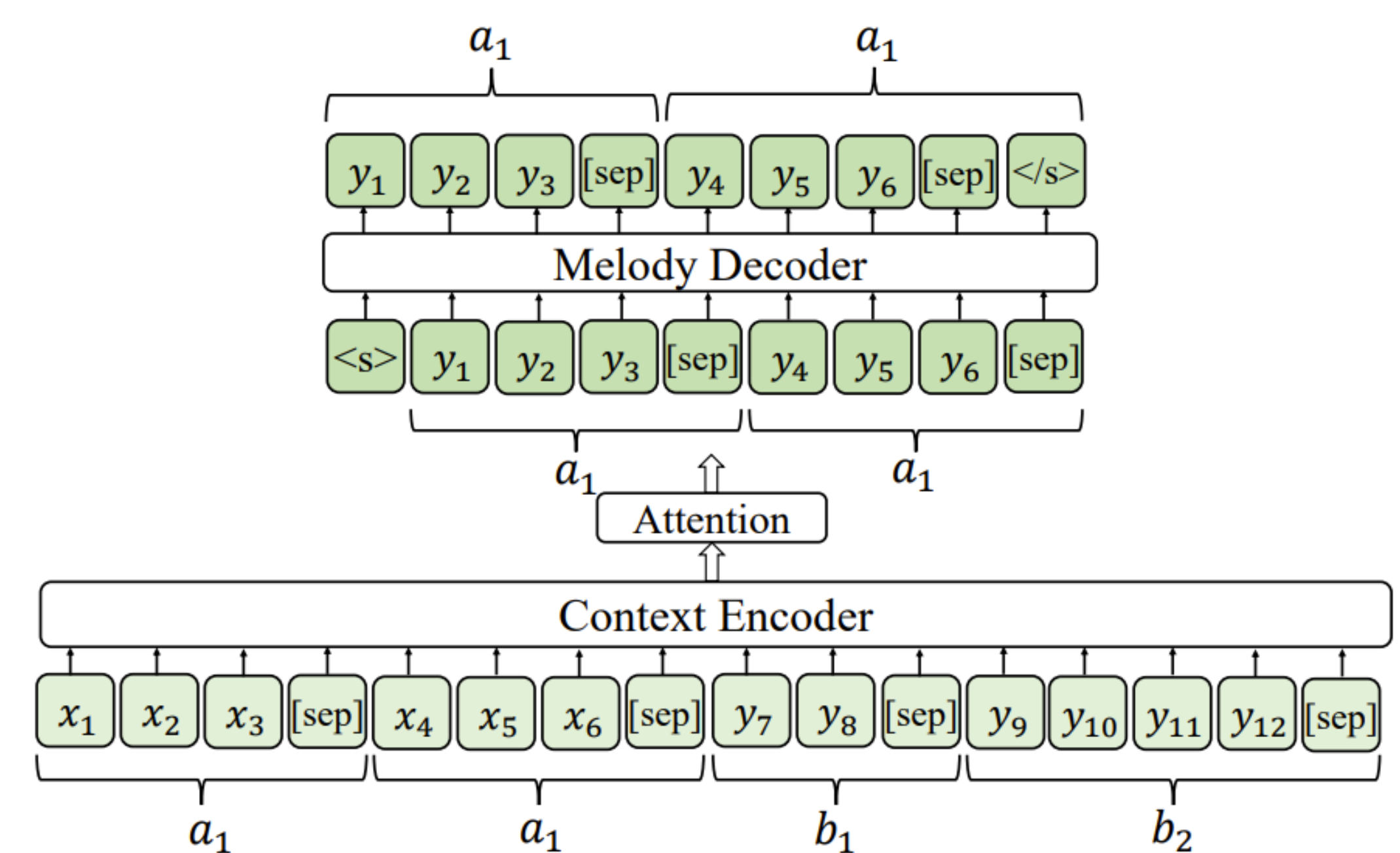
System Architecture



(a) Pipeline of MeloForm.



(b) Detailed process of the expert system.



(c) Architecture of melody refinement neural networks. The input contains four phrases from a melody with musical form as $A(a_1, a_1)B(b_1, b_2)$. In each phrase, x represents the conditioning information (i.e., rhythm, chord and cadence) for each note, while y represents the melody information (i.e., rhythm and pitch) for each note. [sep] indicates the boundaries between phrases, while <s> and </s> indicates the beginning and end of phrases.

Conclusions

1. In this paper, we propose MeloForm, a system to generate melody with musical form based on expert systems and neural networks.
2. It combines the advantages from expert systems to precisely control musical form and neural networks to refine melody for better musical richness without changing musical form.
3. Experimental results demonstrate MeloForm achieves 97.79% accuracy in musical form control, and outperforms baseline systems in structure, thematic, richness and overall quality in terms of subjective evaluation.

Related Link

Demo page: <https://ai-music.github.io/meloform/>

Code link: <https://github.com/microsoft/muzic>