Modeling Melodic Feature Dependency with Modularized Variational Auto-Encoder

Yu-An Wang* Yu-Kai Huang* Tzu-Chuan Lin* Shang-Yu Su **Yun-Nung (Vivian) Chen**

國立臺灣大學
National Taiwan University

NTU MIULAB

Outline

Motivation

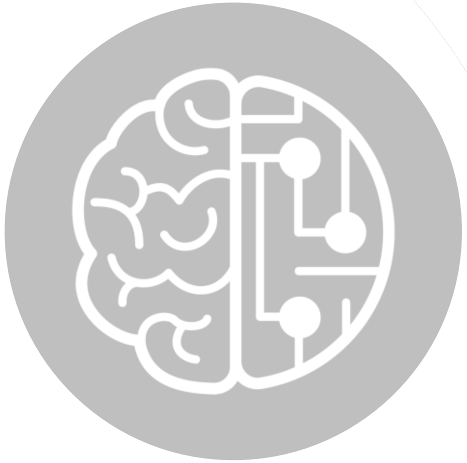
Approach

Experiments

Conclusion

Outline

Motivation



Approach

Experiments

Conclusion

Symbolic Music Generation

Sequence Modeling	Generative Modeling
• RNN (Recurrent Neural Networks)	• VAE (Variational Auto-Encoders) • GAN (Generative Adversarial Networks)

- VRAE = RNN + VAE
  - *modeling temporal dependency via recurrent units*
  - *diverse generation via controllable codes*

Idea: modeling melodic dependency of notes in terms of time, duration, and pitch in a specific order

- Contributions
  - ✓incorporate domain knowledge by a *modularized* framework
  - ✓integrate *note-unrolling* to model the dependency between melodic features
  - ✓achieve better performance than other generative models

Outline

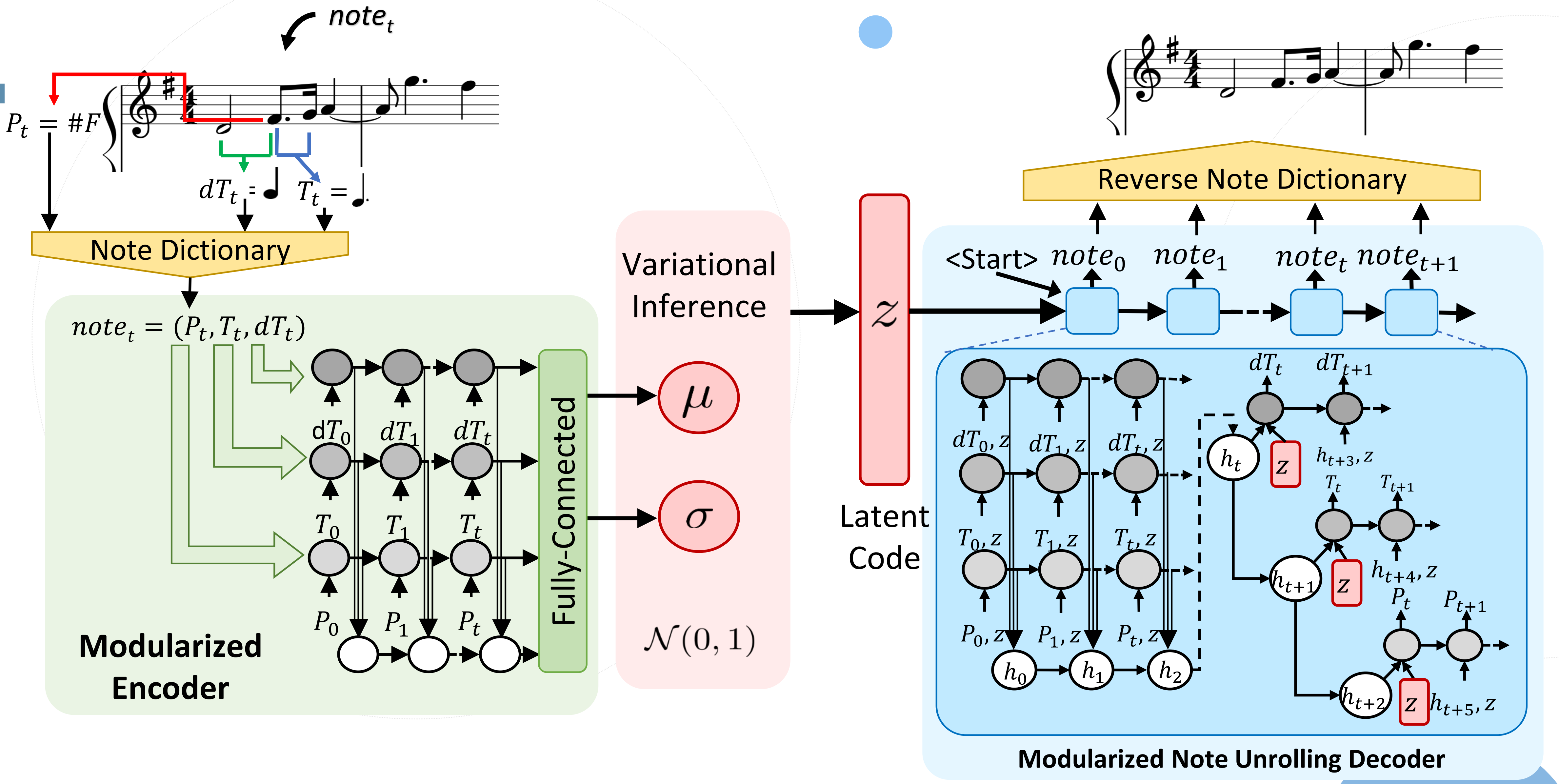
Motivation

Approach

Experiments

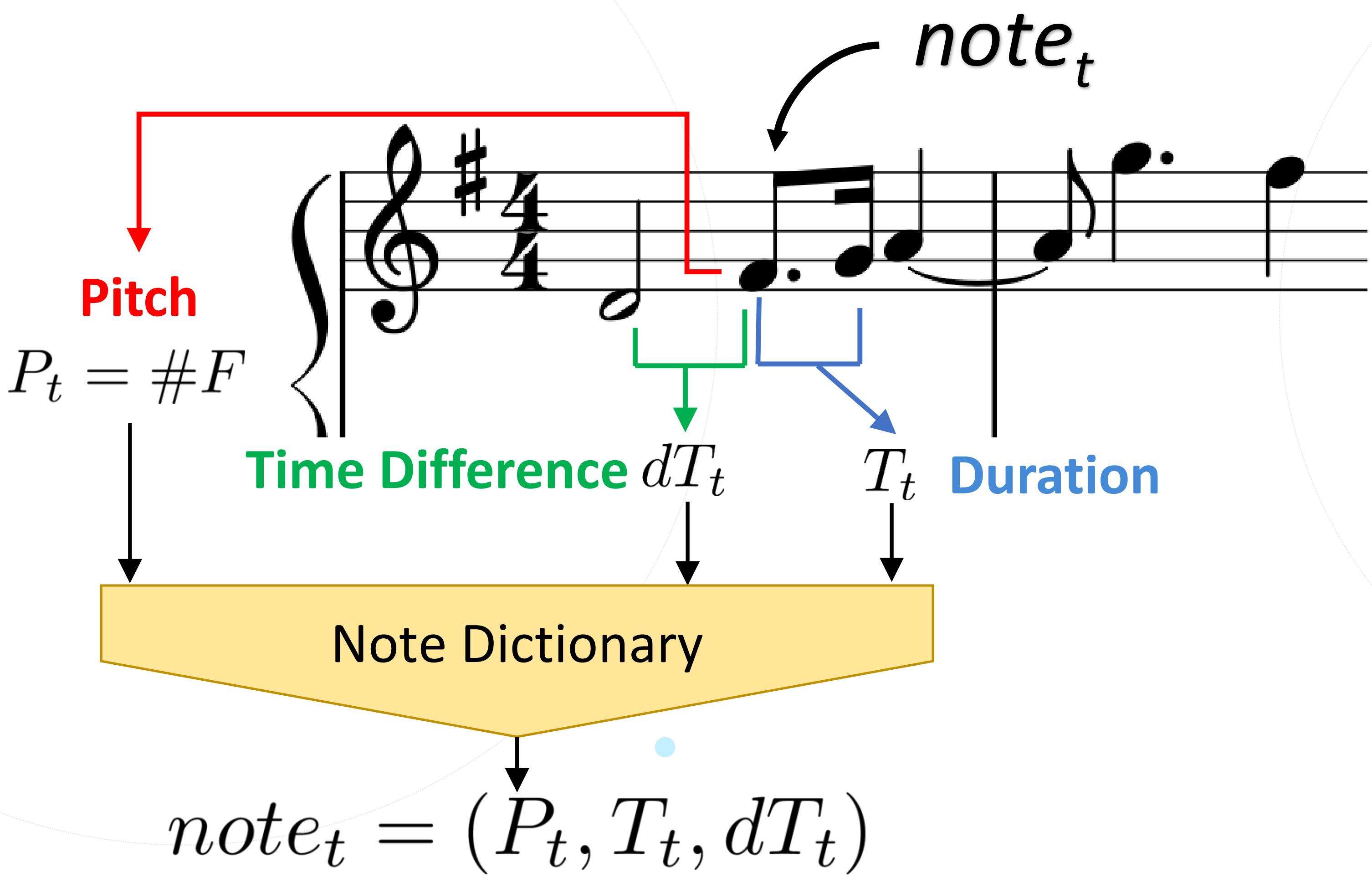
Conclusion

MVAE: Modularized Variational Auto-Encoder



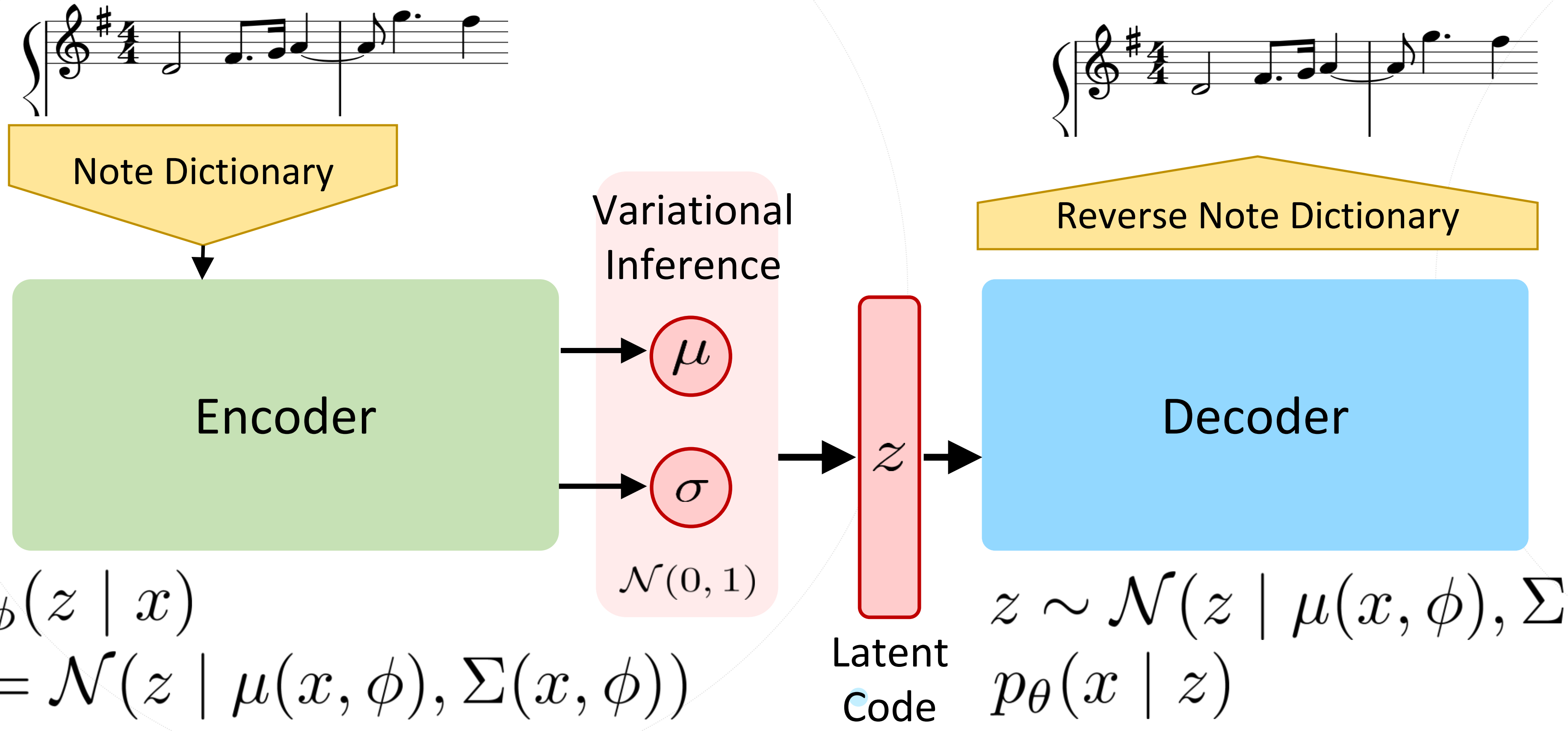
Data Representation

- Represent note events via different features



Variational Auto-Encoder

- Given an input $x = (x_1, \cdots, x_t)$, the goal is to reconstruct the input



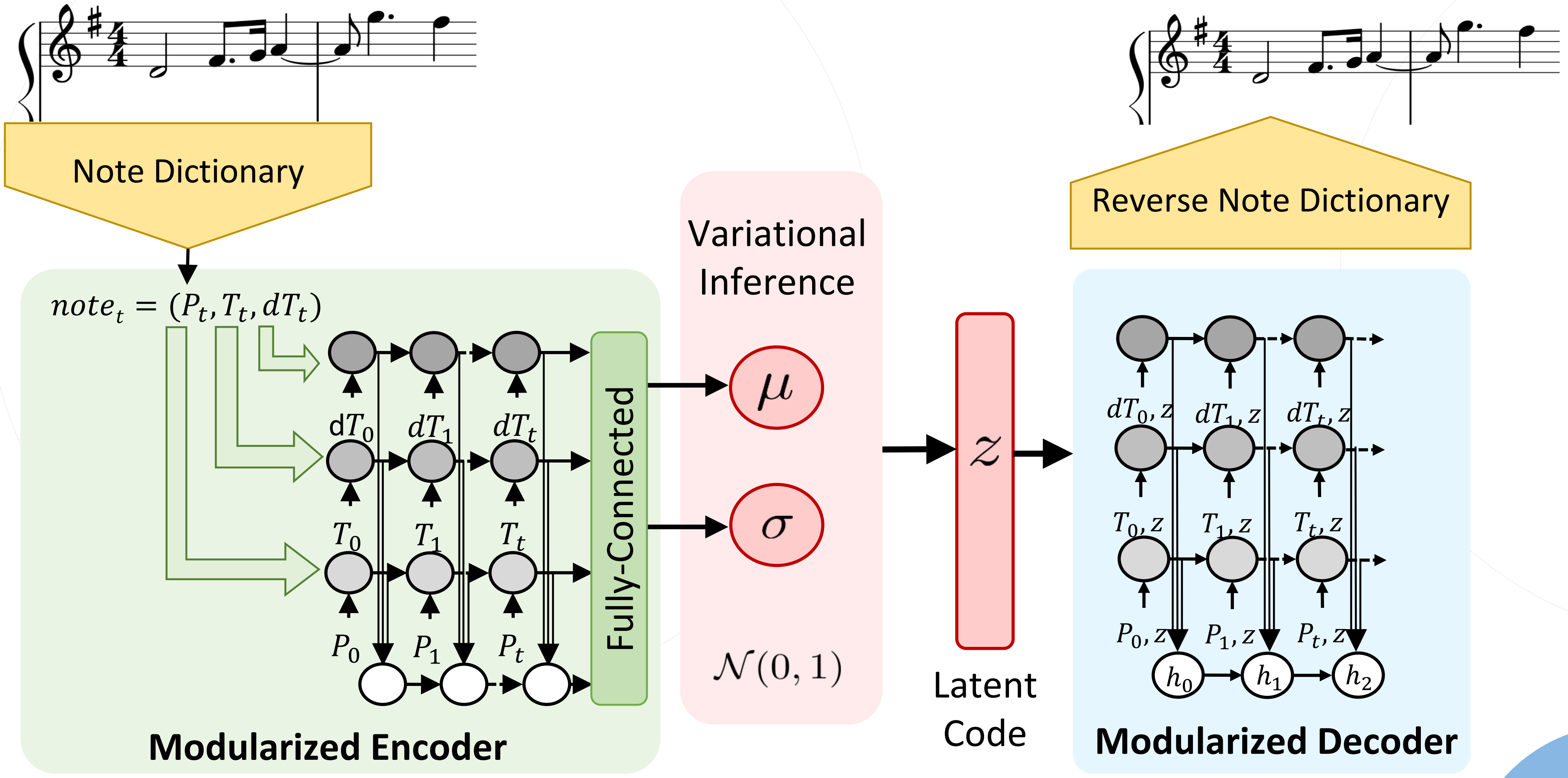
$$q_\phi(z \mid x) = \mathcal{N}(z \mid \mu(x,\phi), \Sigma(x,\phi))$$

$$z \sim \mathcal{N}(z \mid \mu(x,\phi), \Sigma(x,\phi))$$
$$p_\theta(x \mid z)$$

$$\mathcal{L}(x) = \mathrm{E}_{q_\phi(z|x)}[\log p_\theta(x \mid x)] - D_{KL}(q_\phi(z \mid x) \| p_\theta(z))$$

Modularized Encoder and Decoder

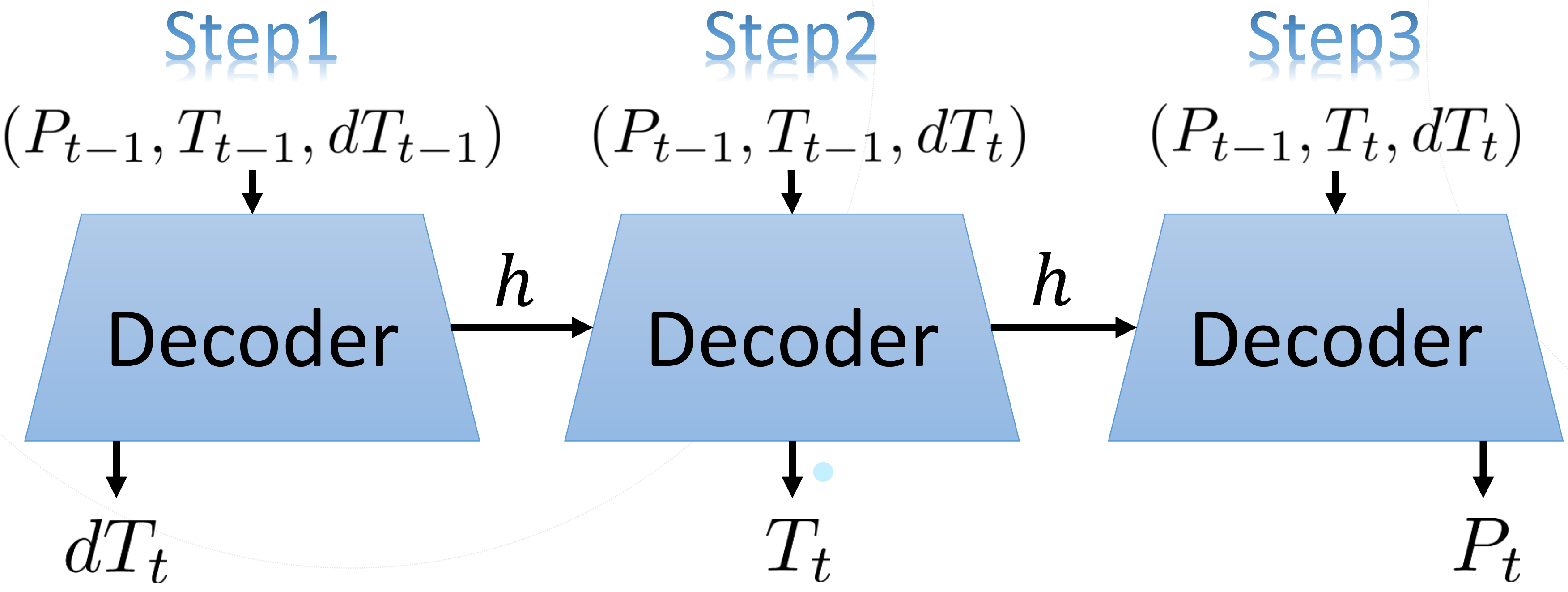
- Each feature is modeled by its own RNN



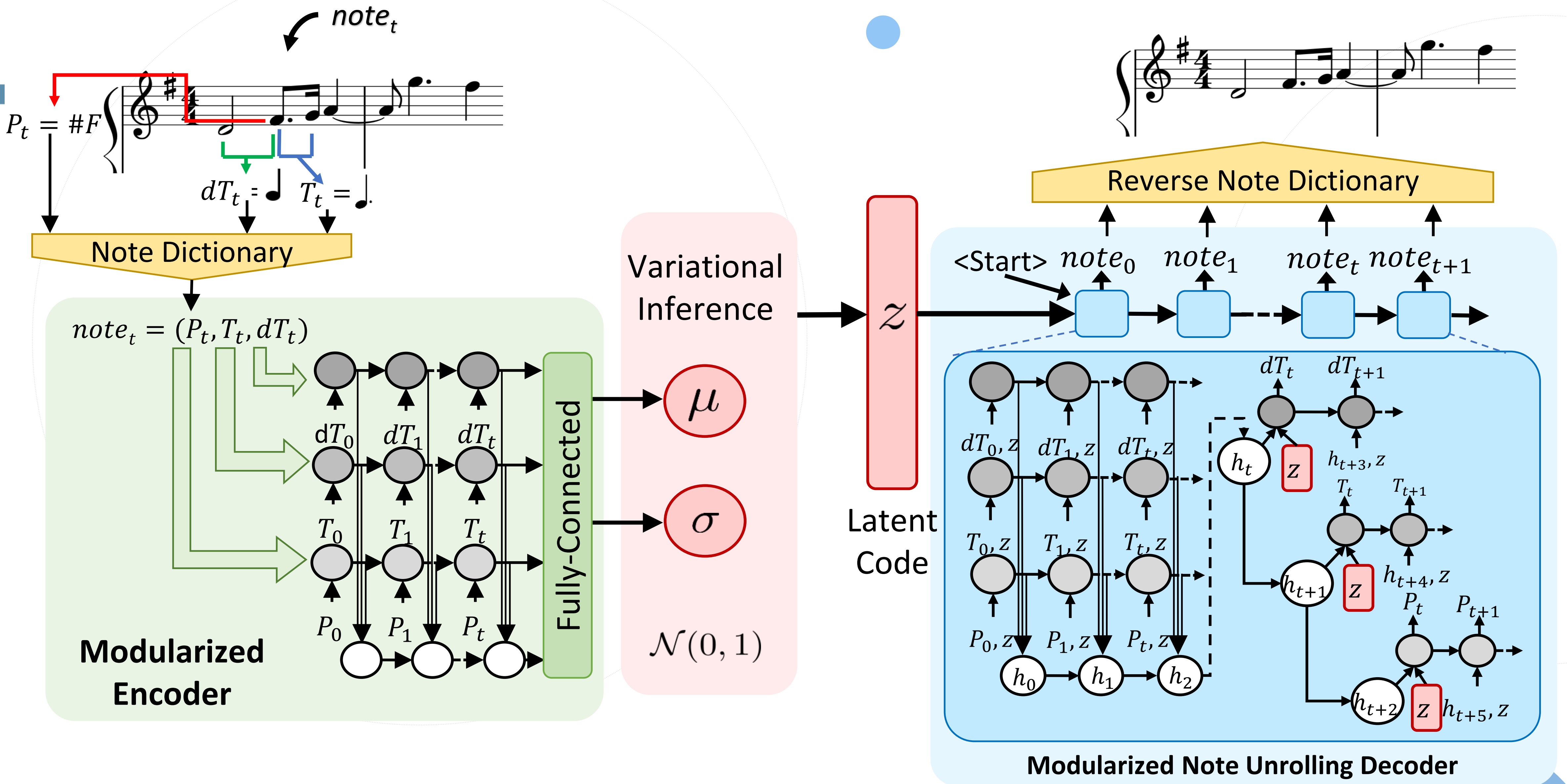
Modularized Note-Unrolling Decoder

- Modeling inter-feature dependency in a specific order

$$p(P_t, T_t, dT_t \mid \text{note}_{1:t-1}) \\ = p(dT_t \mid \text{note}_{1:t-1}) \times p(T_t \mid \text{note}_{1:t-1}) \times p(P_t \mid \text{note}_{1:t-1})$$

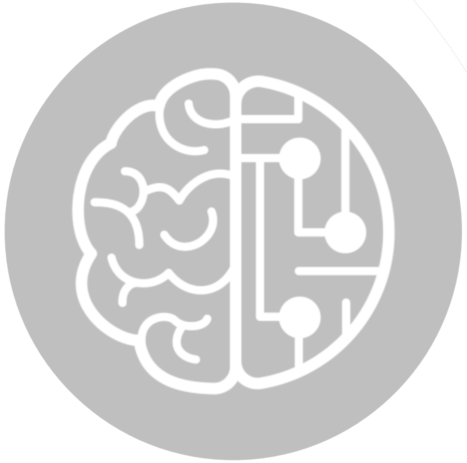


MVAE: Modularized Variational Auto-Encoder



Outline

Motivation



Approach

Experiments

Conclusion

Experimental Setup

- Data: a merged set of Nottingham, Piano-midi.de, JSB Chorales

- Q1: Is the *modularized encoder* better?
  - Baseline – BachProp (Colombo & Gerstner, 2018)
- Q2: Is the *variational* inference important?
  - Baseline – modularized auto-encoder
- Q3: Is the *note-enrolling* important?
  - Ablation test

Florian Colombo and Wulfram Gerstner, “Bachprop: Learning to compose music in multiple styles,” CoRR, 2018.

Human Evaluation

- 1-6 scales (1: machine-generated; 6: human-generated)
- Collect 85 scores for each model

Model	Reconstruction Error	KL Divergence	Human Score	
			μ	σ
BachProp	240.16	-	3.51	1.61
Modularized AutoEncoder	20.79	-	2.77	1.65
Proposed w/o note unrolling	85.88	264.00	3.22	1.73
Proposed w/ note unrolling	73.19	30.37	4.24	1.54
Real data	-	-	4.34	1.55

✓ A1: The *modularized encoder* is better.
✓ A2: The *variational* inference is necessary.
✓ A3: The *note-enrolling* is important.

Latent Space Analysis

- Interpolation distribution

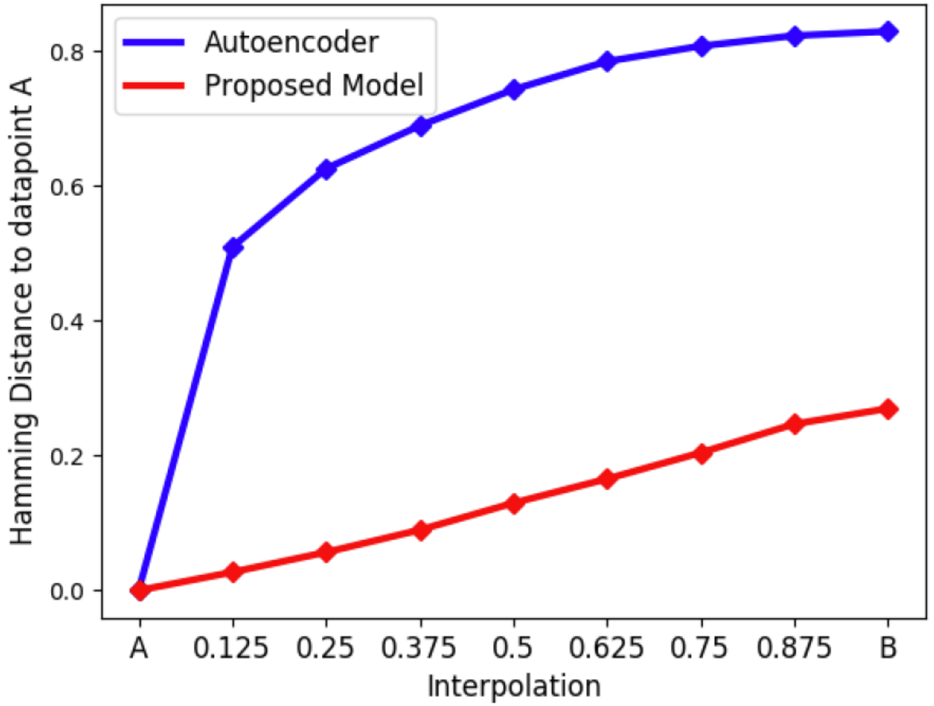


Fig. 2. Average distance between two random datapoints on Z.

Smooth curve: meaningful interpolation points

- Visualization

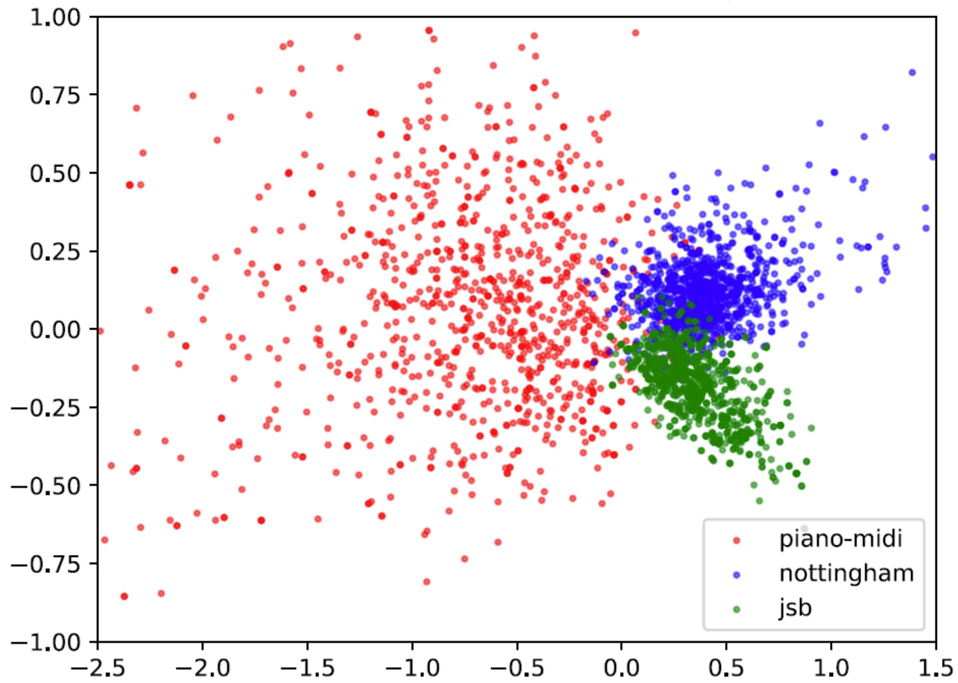
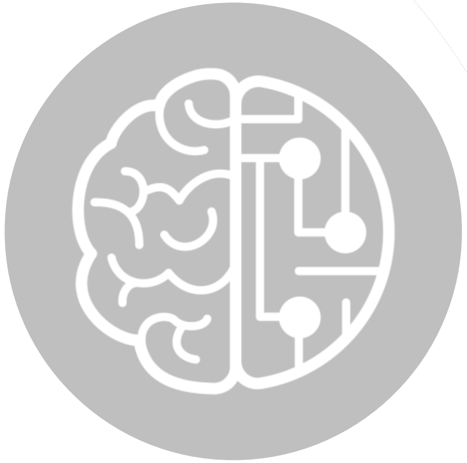


Fig. 3. Visualization on the latent space via PCA, where three different types of music are separated in Z.

Distinct features for different music characteristics
→ informative latent codes

Outline

Motivation



Approach

Experiments

Conclusion

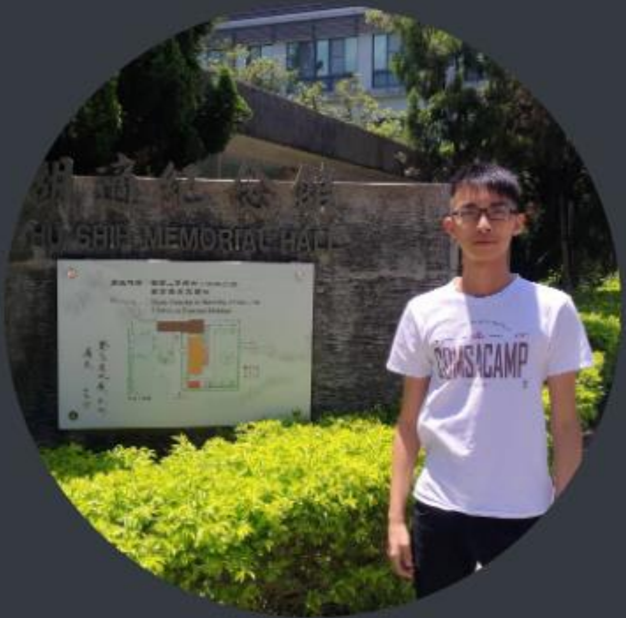
Conclusion

- We propose a VAE with a *modularized* framework to model the melodic dependency between note attributes
- The proposed note event representations bring better flexibility
- The experiments in a merged dataset with diverse music types show the superior performance of our MVAE
  - ✓ The *modularized encoder* is better.
  - ✓ The *variational* inference is necessary.
  - ✓ The *note-enrolling* is important.
  - ✓ The learned latent codes are informative

NTU MIULAB

Yu-An Wang*

Yu-Kai Huang*



Tzu-Chuan Lin*

MiuLaber

Shang-Yu Su

Yun-Nung Chen

Demo & Code Available @

- http://mvae.miulab.tw