Adaptive Accompaniment with **ReaLchords**

Yusong Wu, Tim Cooijmans, Kyle Kastner, Adam Roberts, Ian Simon, Alexander Scarlatos, Chris Donahue, Cassie Tarakajian, Shayegan Omidshafiei, Aaron Courville, Pablo Samuel Castro, Natasha Jaques, Cheng-Zhi Anna Huang

Université de Montréal

Mila

Google DeepMind

LLM & Generative AIs are Great...

Gemini

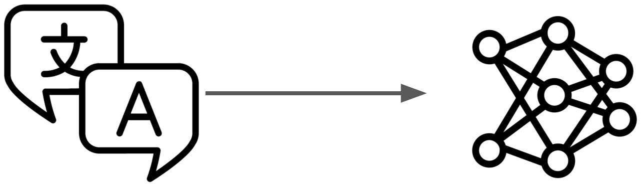
Stable Audio

LLM & Generative AIs are Great, **but:**

Designed for “***offline***” usage:

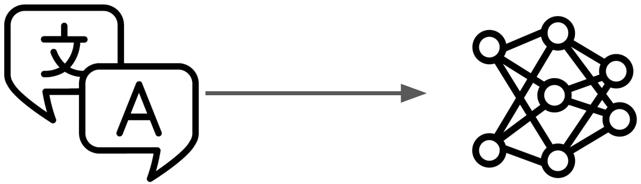
LLM & Generative AIs are Great, **but:**

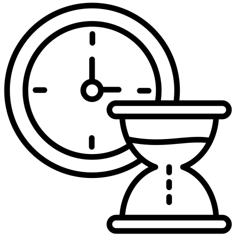
Designed for “***offline***” usage:



LLM & Generative AIs are Great, **but:**

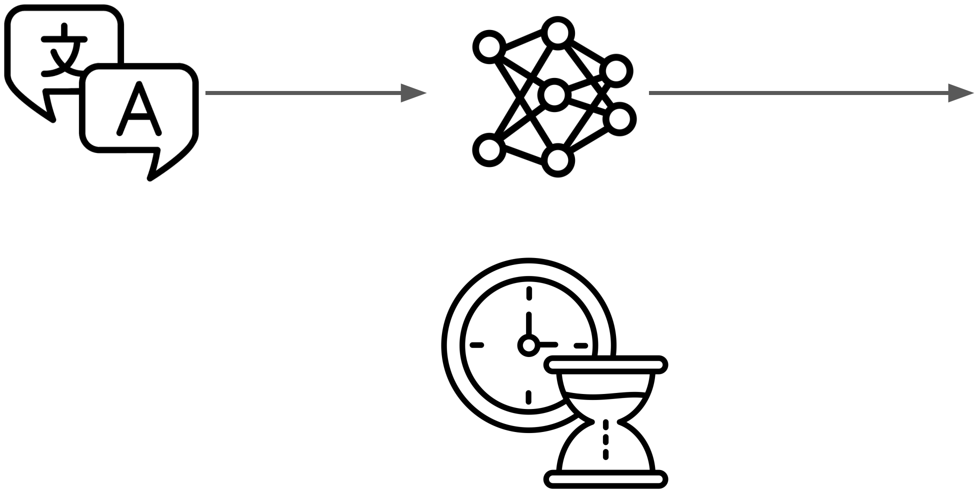
Designed for “***offline***” usage:

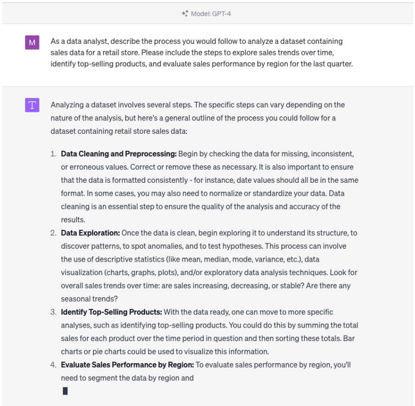




LLM & Generative AIs are Great, **but:**

Designed for "***offline***" usage:





Model: GPT-4

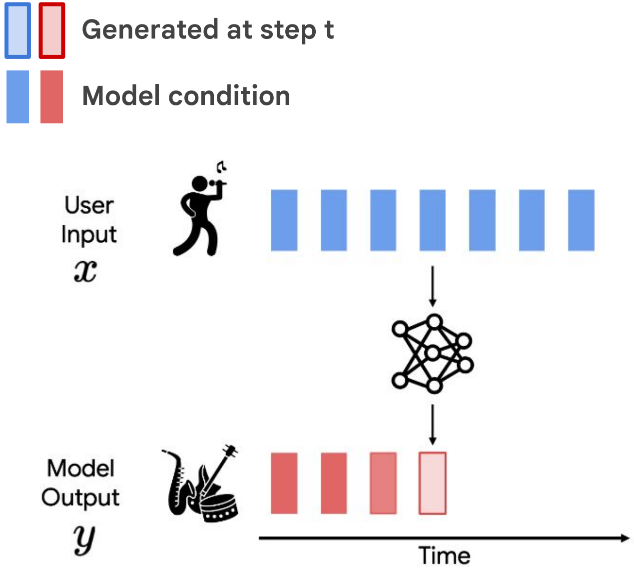
M As a data analyst, describe the process you would follow to analyze a dataset containing sales data for a retail store. Please include the steps to explore sales trends over time, identify top-selling products, and evaluate sales performance by region for the last quarter.

T Analyzing a dataset involves several steps. The specific steps can vary depending on the nature of the analysis, but here's a general outline of the process you could follow for a dataset containing retail store sales data:

1. **Data Cleaning and Preprocessing:** Begin by checking the data for missing, inconsistent, or erroneous values. Correct or remove these as necessary. It is also important to ensure that the data is formatted consistently - for instance, date values should all be in the same format. In some cases, you may also need to normalize or standardize your data. Data cleaning is an essential step to ensure the quality of the analysis and accuracy of the results.
2. **Data Exploration:** Once the data is clean, begin exploring it to understand its structure, to discover patterns, to spot anomalies, and to test hypotheses. This process can involve the use of descriptive statistics (like mean, median, mode, variance, etc.), data visualization (charts, graphs, plots), and/or exploratory data analysis techniques. Look for overall sales trends over time: are sales increasing, decreasing, or stable? Are there any seasonal trends?
3. **Identify Top-Selling Products:** With the data ready, one can move to more specific analyses, such as identifying top-selling products. You could do this by summing the total sales for each product over the time period in question and then sorting these totals. Bar charts or pie charts could be used to visualize this information.
4. **Evaluate Sales Performance by Region:** To evaluate sales performance by region, you'll need to segment the data by region and

Human-AI Interaction: **Simultaneous, Real-time**

Generative Models are **NOT** for **Live Interaction**



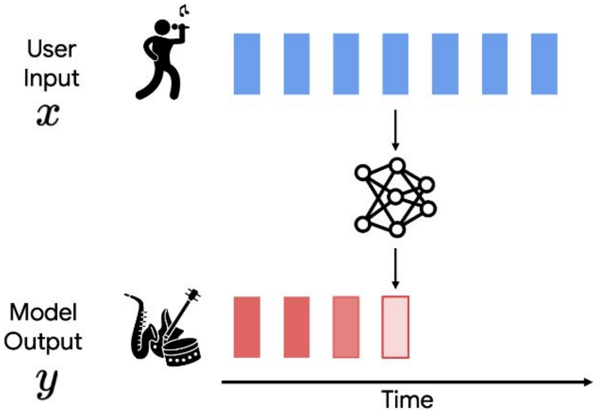
Offline Models

$$\phi(y_t | x, y_{<t})$$

Generative Models are **NOT** for **Live Interaction**

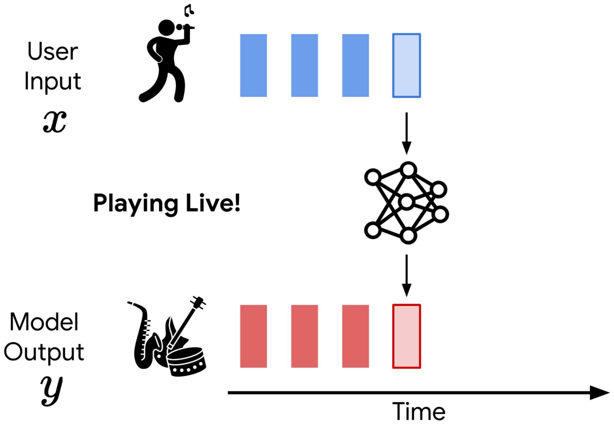
Generated at step t

Model condition



Offline Models

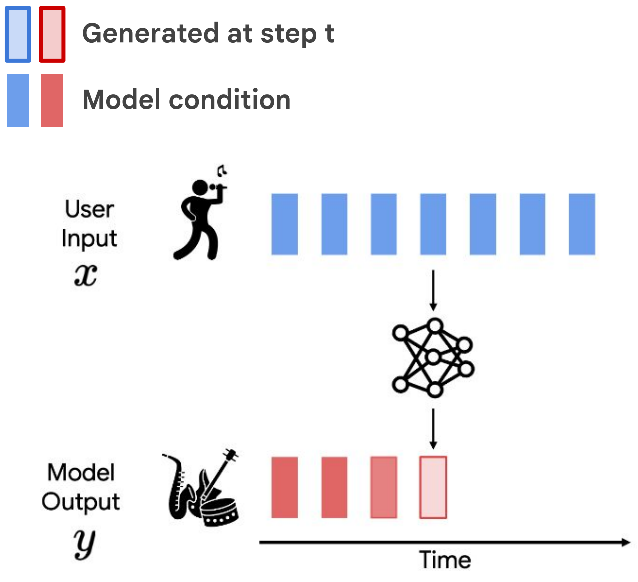
$$\phi(y_t | x, y_{<t})$$



Online Models

$$\pi_\theta(y_t | x_{<t}, y_{<t})$$

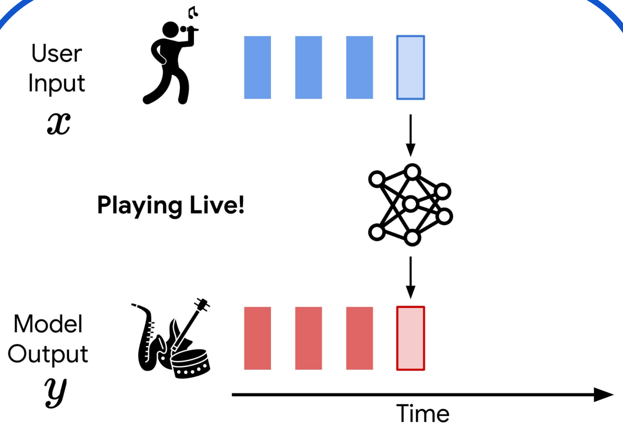
Generative Models are NOT for Live Interaction



Offline Models

$$\phi(y_t | x, y_{<t})$$

Ideal for Live Interaction



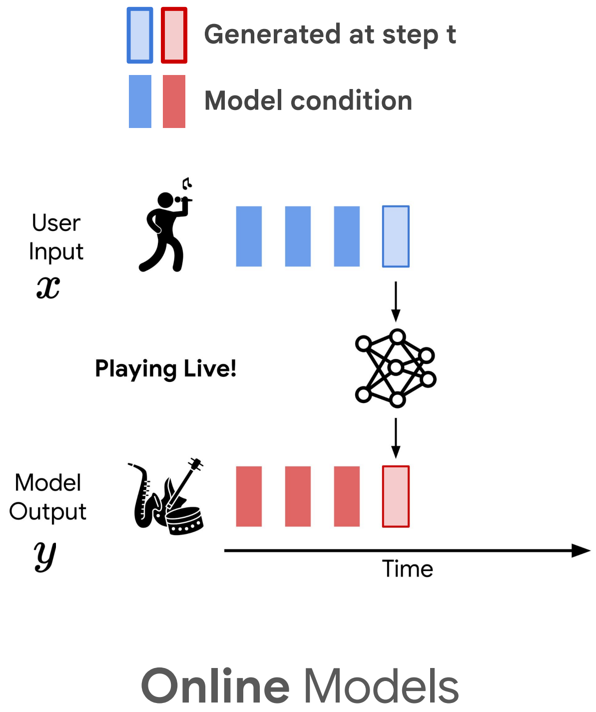
Online Models

$$\pi_\theta(y_t | x_{<t}, y_{<t})$$

Online Models

- Generate **simultaneously** with input:

$$y \sim \prod_{t}^{T} \pi_\theta(y_t \mid x_{<t}, y_{<t})$$

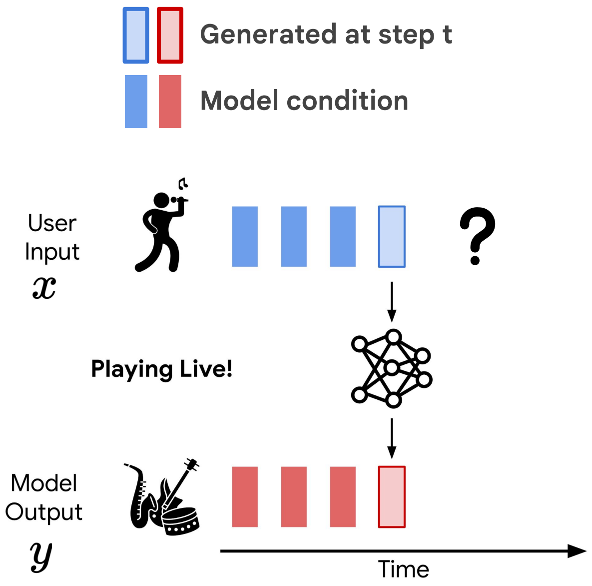


Online Models

- Generate **simultaneously** with input:

$$y \sim \prod_{t}^{T} \pi_\theta(y_t \mid x_{<t}, y_{<t})$$

- Challenge: **“future” unavailable** when generate “present”



Online Models

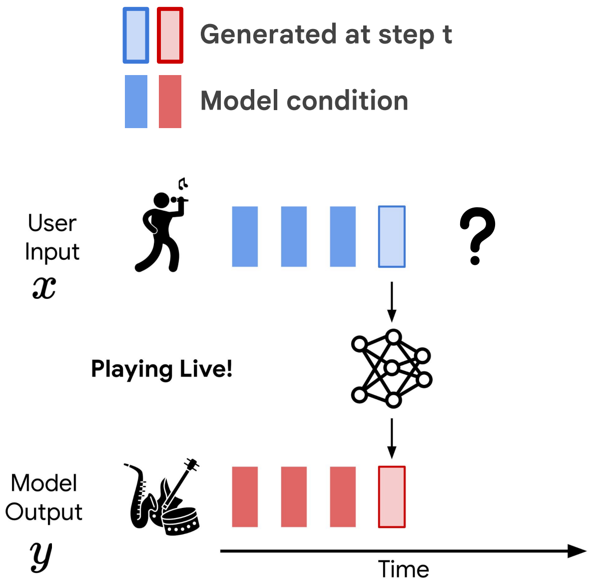
Online Models

- Generate **simultaneously** with input:

$$y \sim \prod_{t}^{T} \pi_\theta(y_t \mid x_{<t}, y_{<t})$$

- Challenge: **“future” unavailable** when generate “present”

  $\longrightarrow$ Must **anticipate the input**



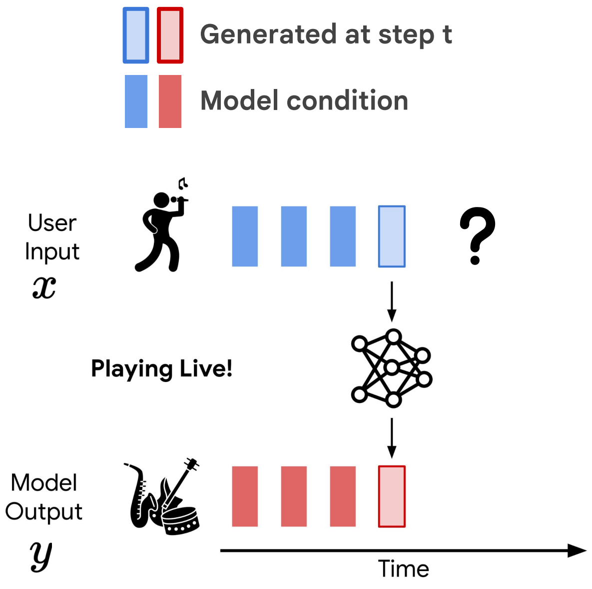
Online Models

Online Models

- Generate **simultaneously** with input:

$$y \sim \prod_{t}^{T} \pi_\theta(y_t \mid x_{<t}, y_{<t})$$

- Challenge: **“future” unavailable** when generate “present”
  - → Must **anticipate the input**
  - → When misanticipate, must **recover from error**



Online Models

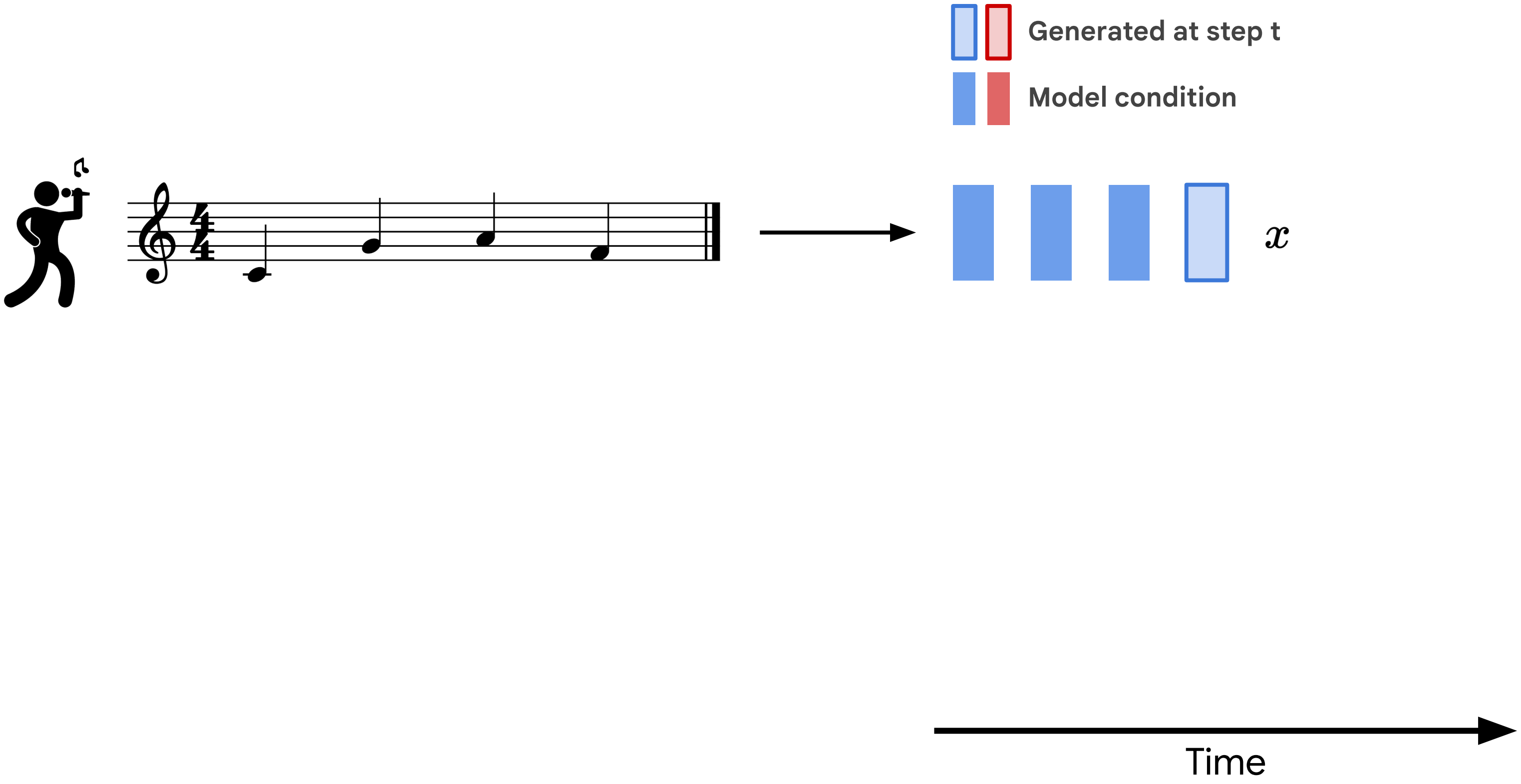
Music Jamming

Simultanous interplay of **coordination, anticipation, and collaborative creativity**

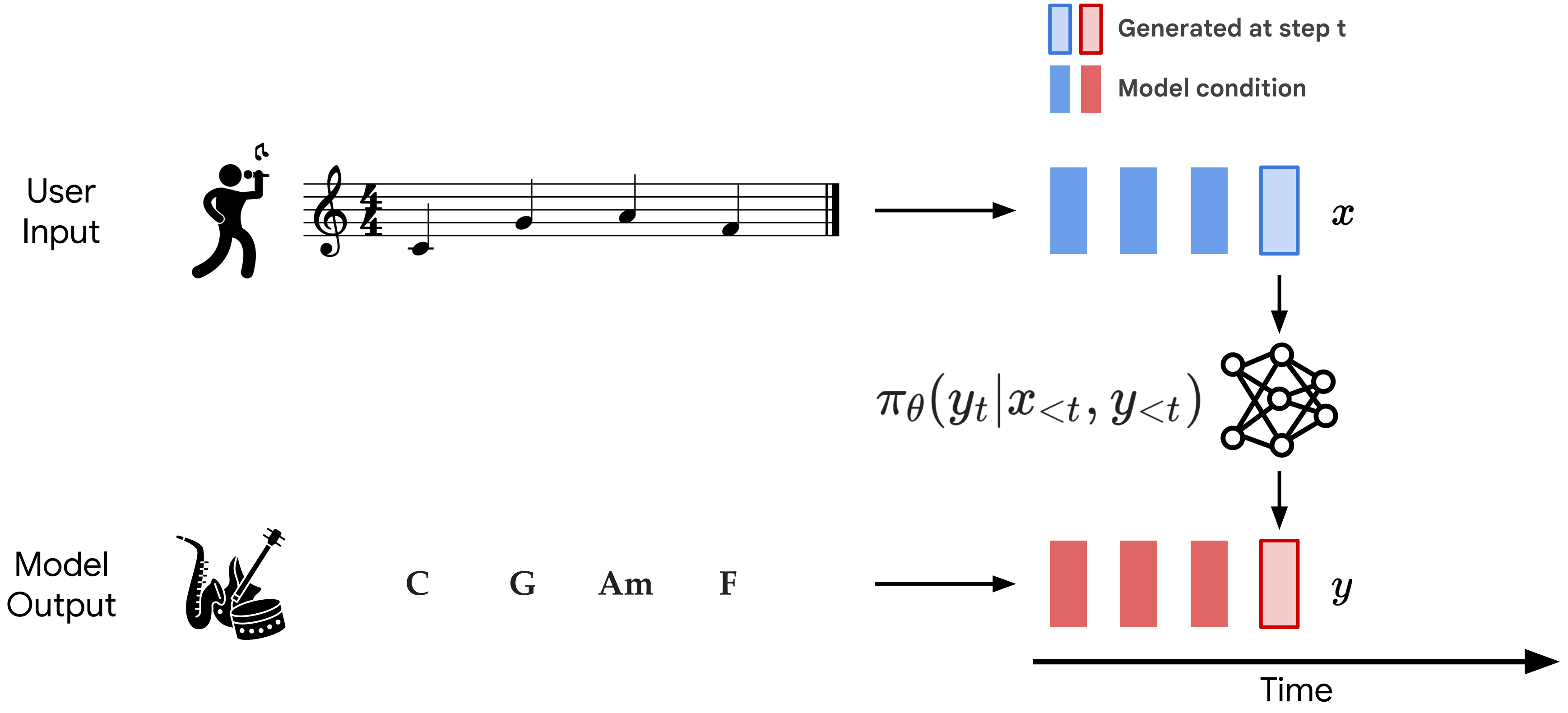
https://www.flickr.com/photos/gutweincreative/9975000744/

Simultaneous Melody-to-Chord Accompaniment

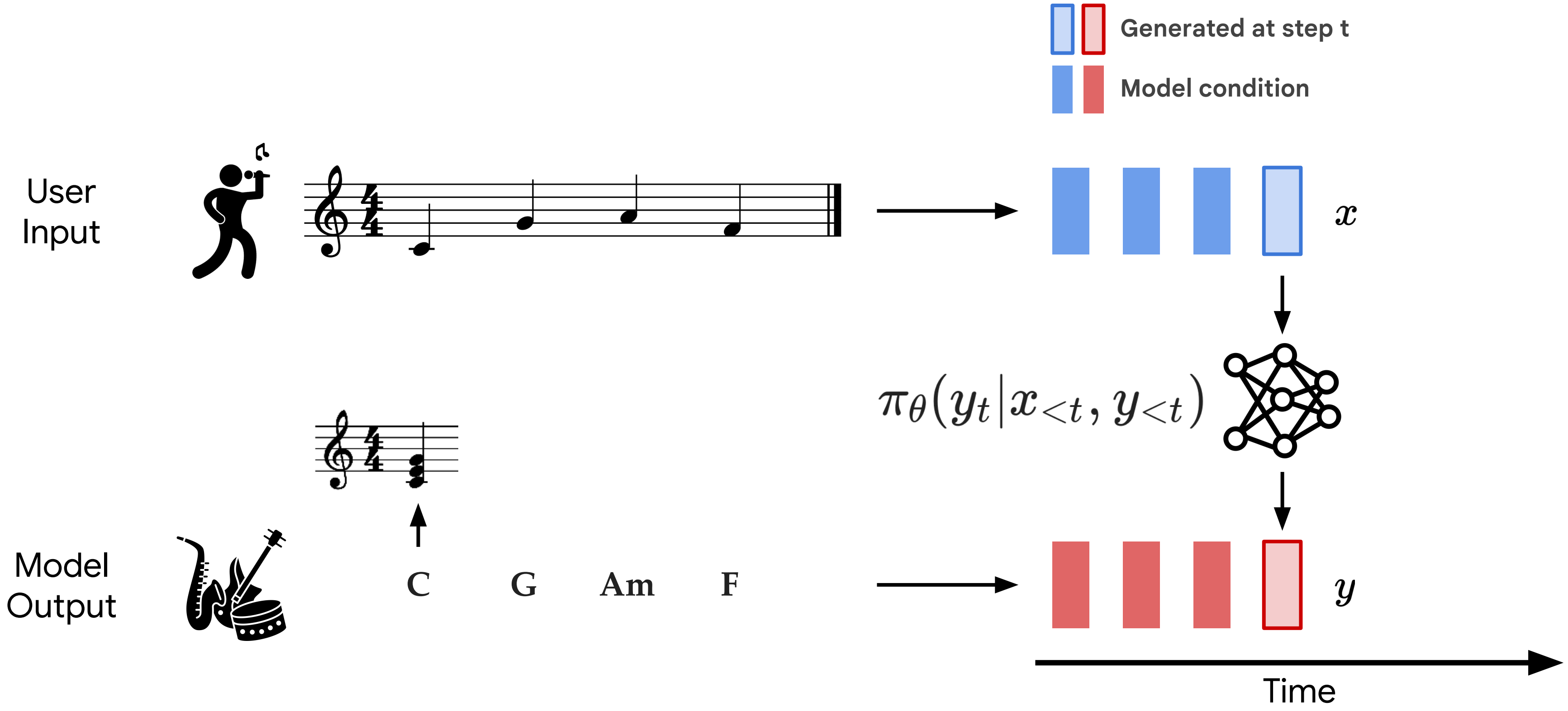
User
Input



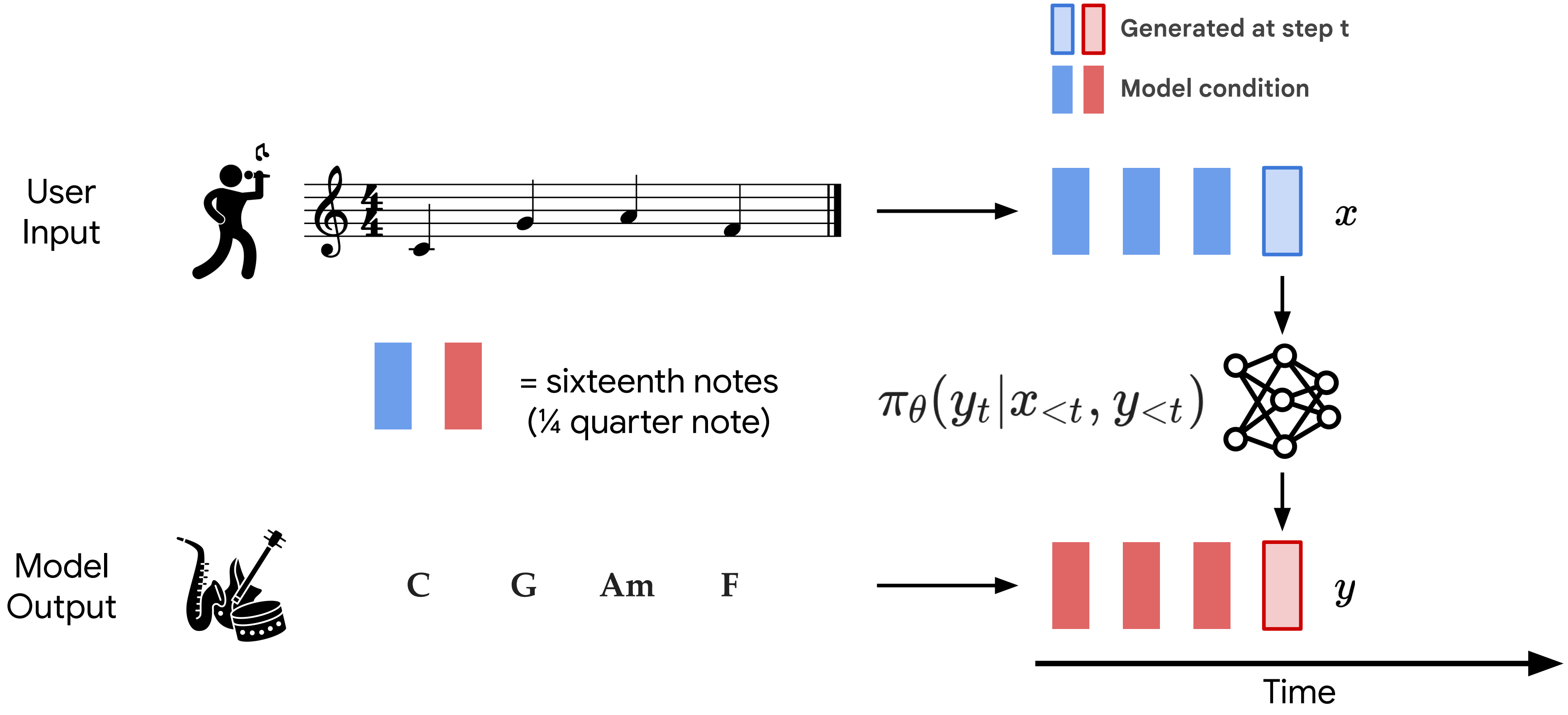
Simultaneous Melody-to-Chord Accompaniment



Simultaneous Melody-to-Chord Accompaniment

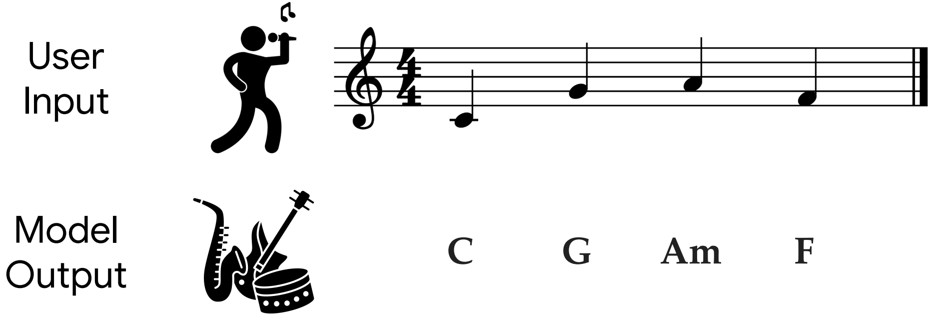


Simultaneous Melody-to-Chord Accompaniment



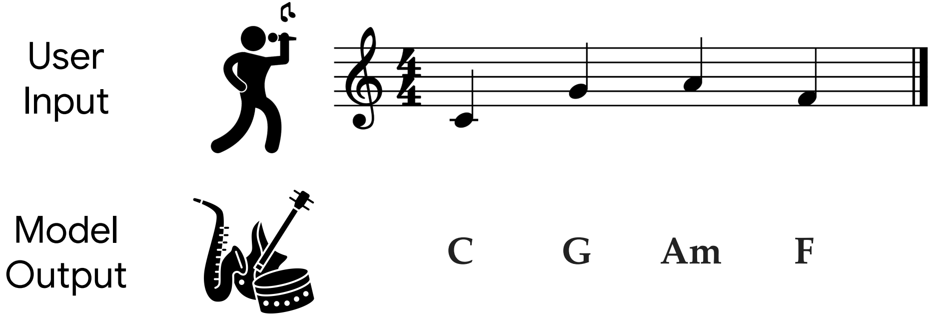
Model and Dataset

- Dataset: hooktheory dataset
  - 38k melody-chord pairs of pop song
  - No human preference label



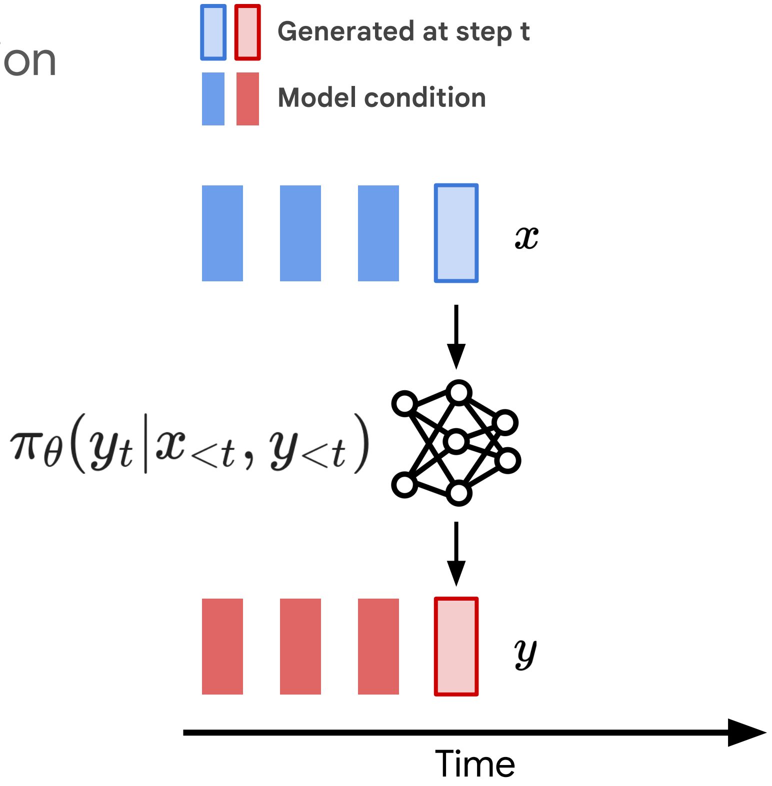
Model and Dataset

- Dataset: hooktheory dataset
  - 38k melody-chord pairs of pop song
  - No human preference label
- Model: 8-layer transformer decoder



Train with MLE: simple but **ineffective**

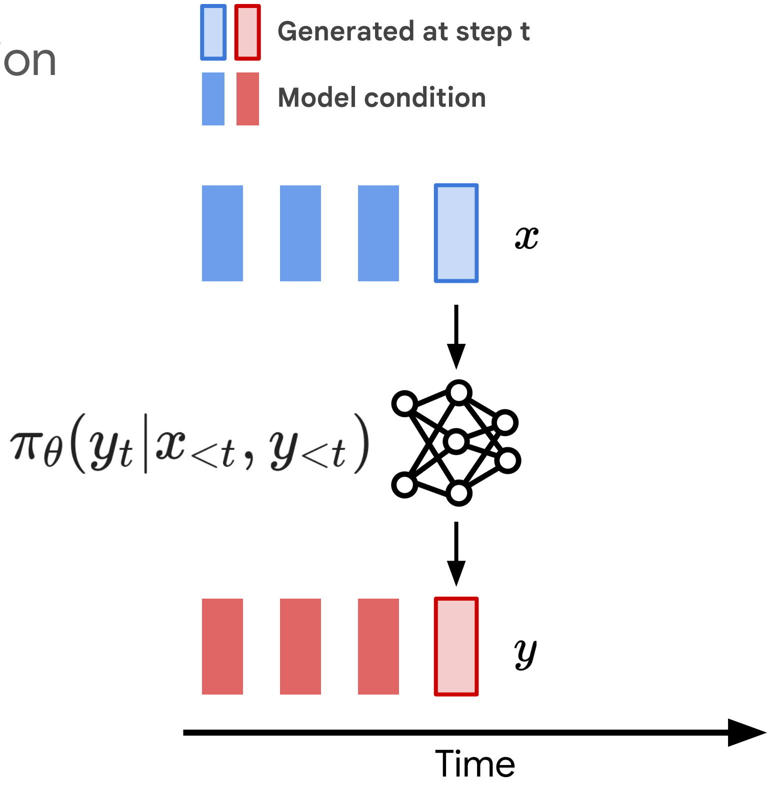
- Training with Maximum Likelihood Estimation (MLE) is straightforward:
    - Next-token prediction + cross entropy



Train with MLE: simple but **ineffective**

- Training with Maximum Likelihood Estimation (MLE) is straightforward:
  - Next-token prediction + cross entropy
- However, MLE will cause exposure bias:
  - Only seen perfect condition from data in training

→ Cannot effectively recover from error



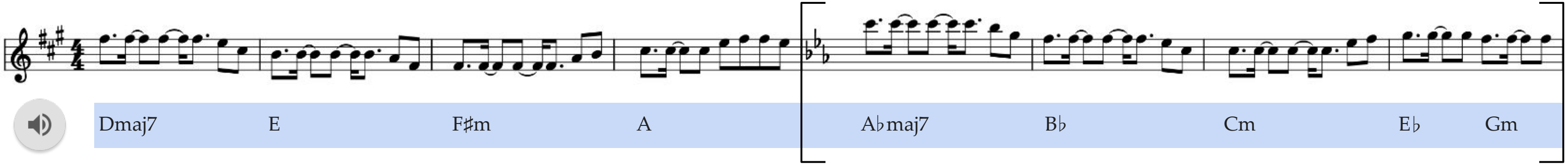
Melody-to-chord accompaniment

Melody-to-chord accompaniment

melody transposed up 6 semi-tones (a tritone)

Input Melody

Ground Truth



Harmonic Chords

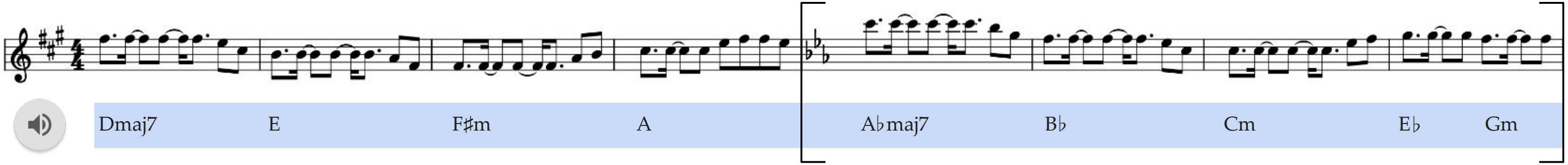
Inharmonic Chords

Melody-to-chord accompaniment

melody transposed up 6 semi-tones (a tritone)

Input Melody

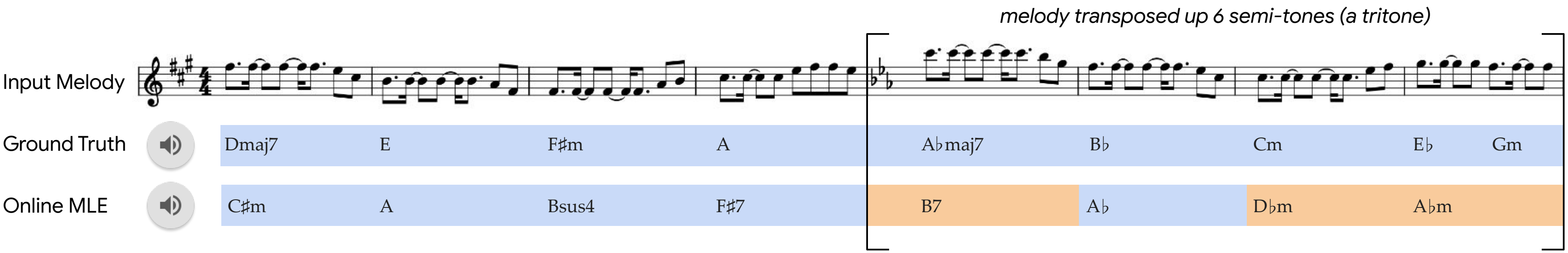
Ground Truth



Harmonic Chords

Inharmonic Chords

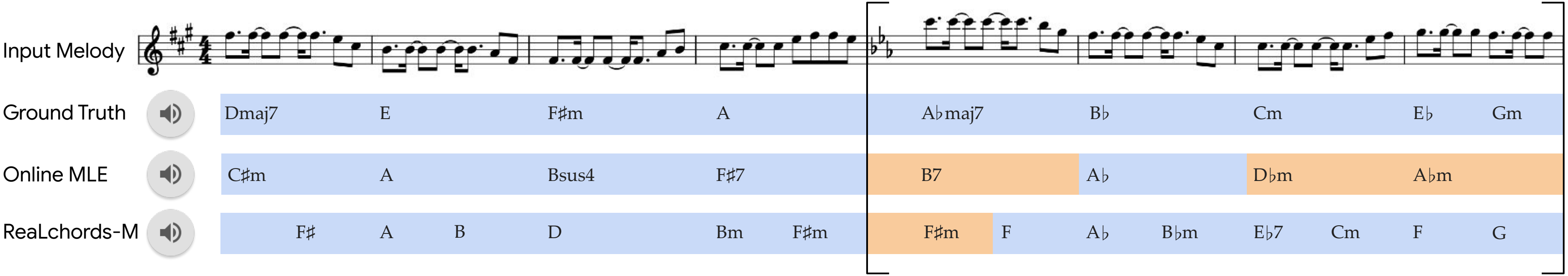
Online generation fails to recover from error



Harmonic Chords | Inharmonic Chords

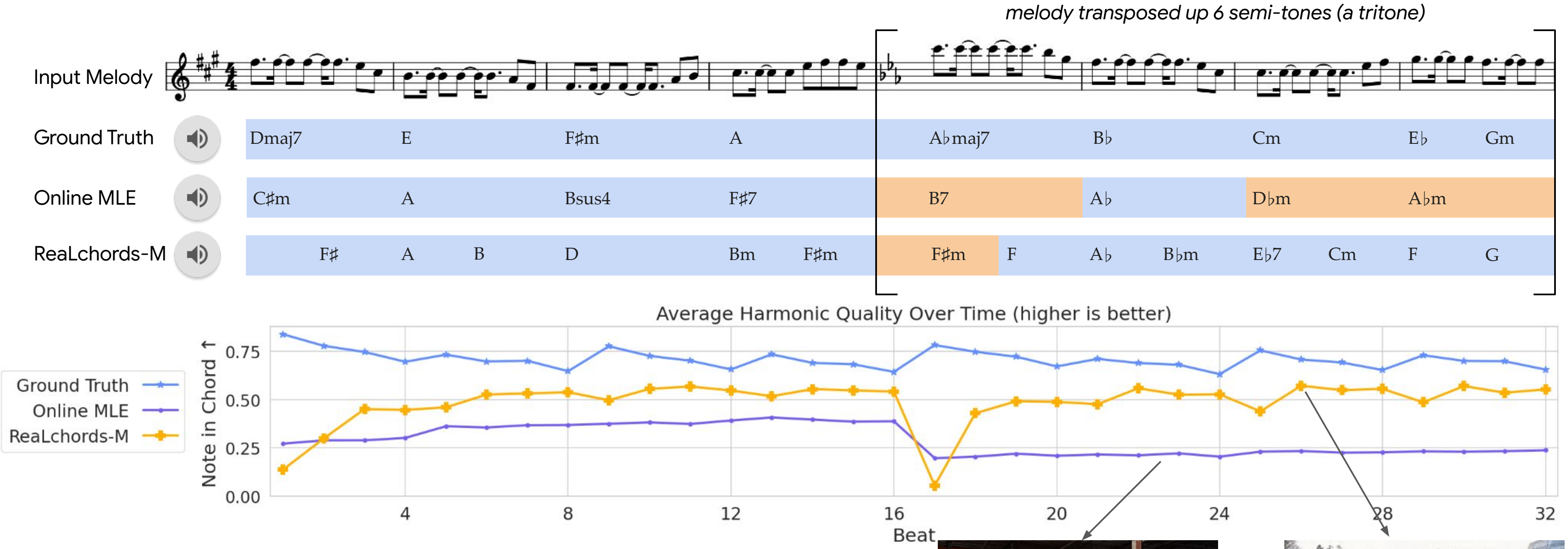
RealChords: adaptive and recovers from error

melody transposed up 6 semi-tones (a tritone)



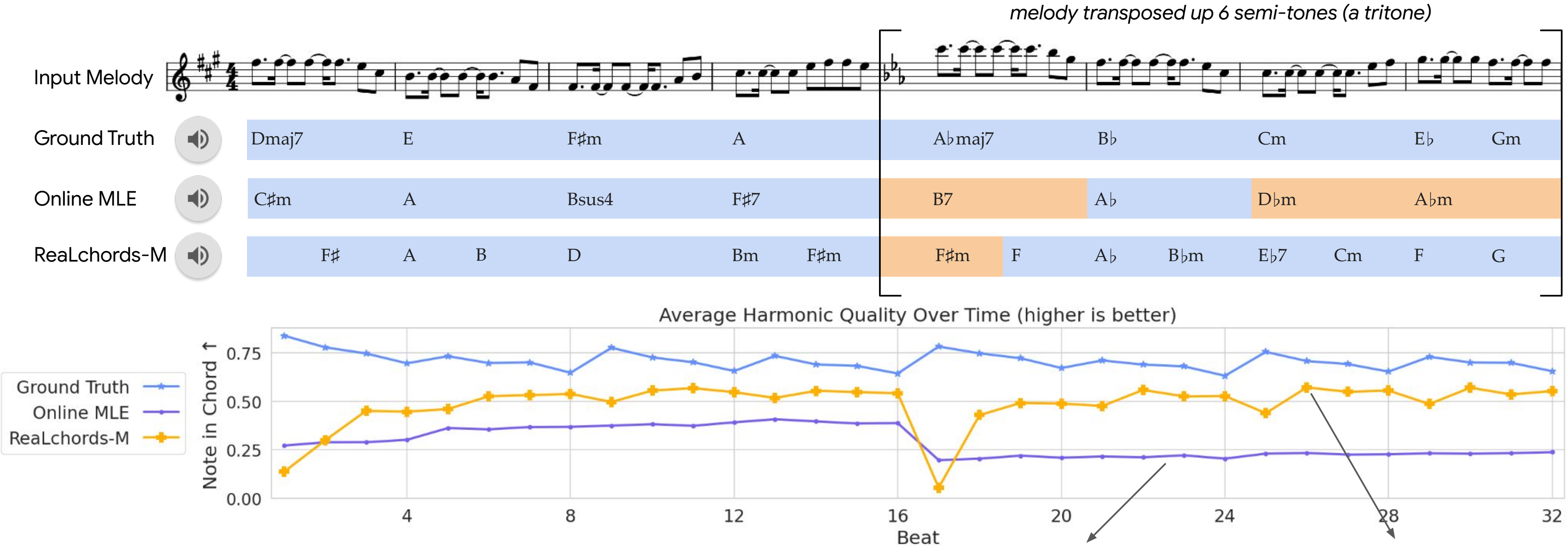
Harmonic Chords | Inharmonic Chords

ReaLchords: adaptive and recovers from error



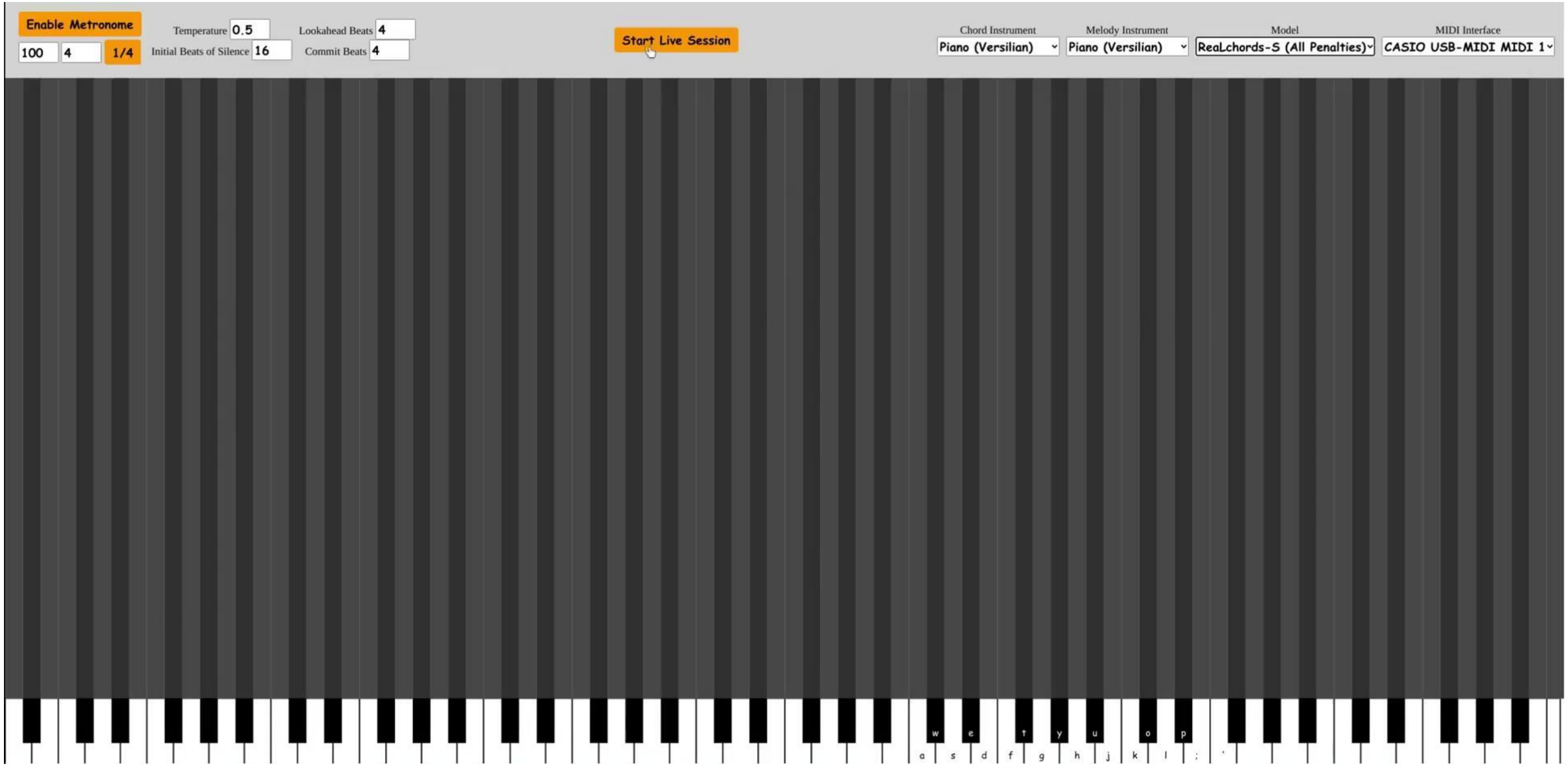
Harmonic Chords **Inharmonic Chords**

ReaLchords: adaptive and recovers from error



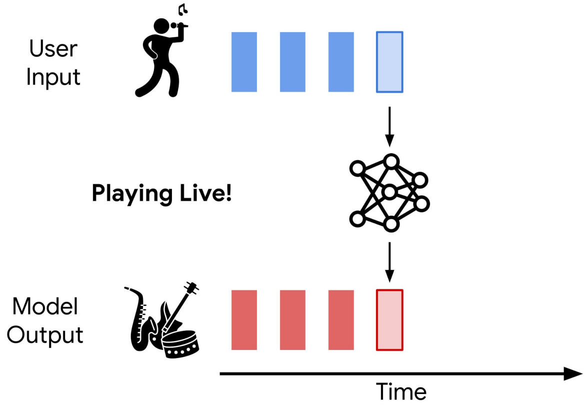
Harmonic Chords   Inharmonic Chords

Powered by **ReaLchords:** Real-time melody-to-chord accompaniment



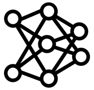
Extra Objectives Needed for Online Model

Extra objectives needed that will tell model how to anticipate and recover from error



Extra Objectives

Extra Objectives

Pretrain from MLE 

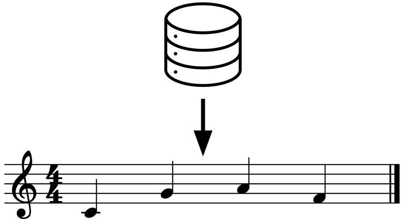
↓

Finetune

Extra Objectives

Pretrain from MLE

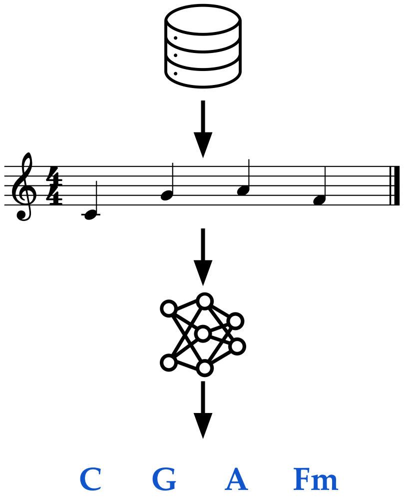
Finetune



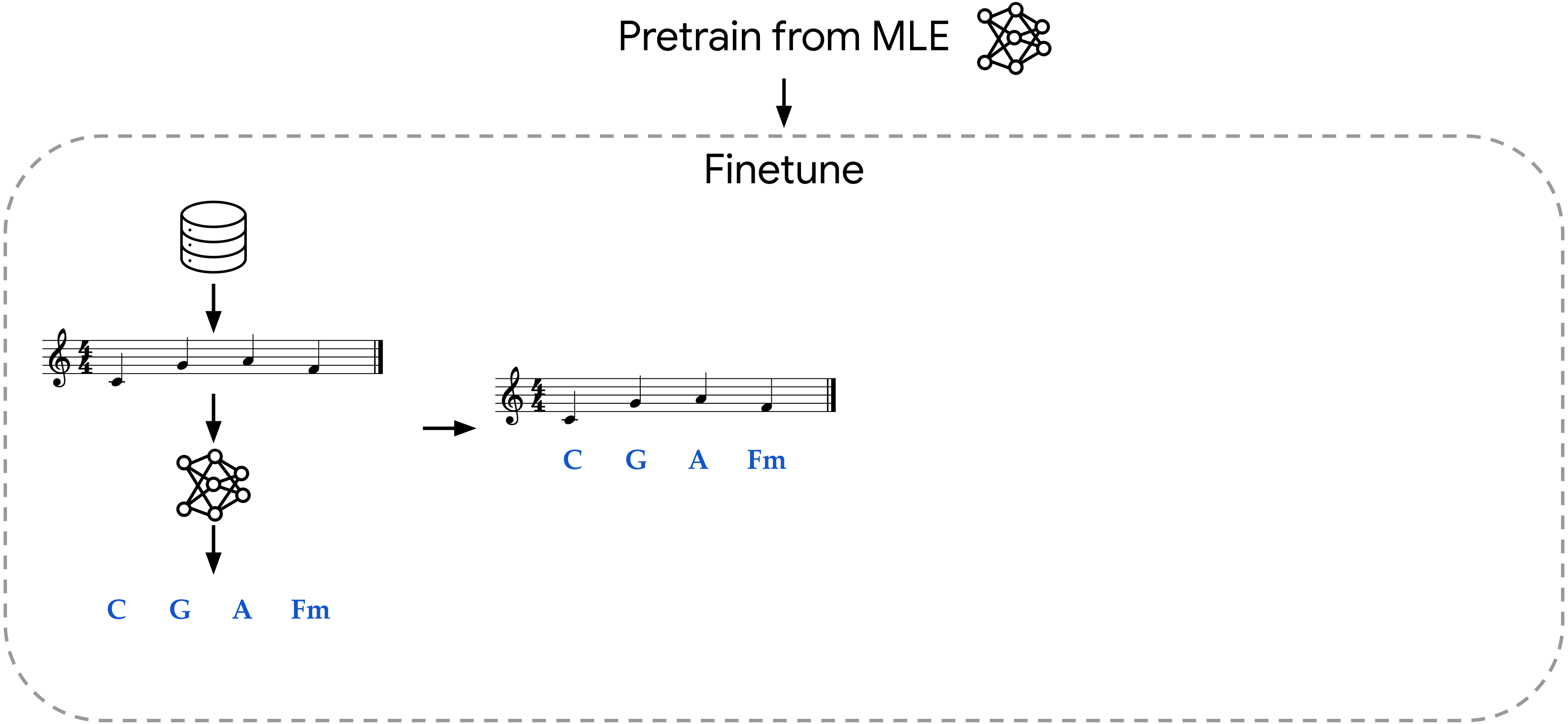
Extra Objectives

Pretrain from MLE

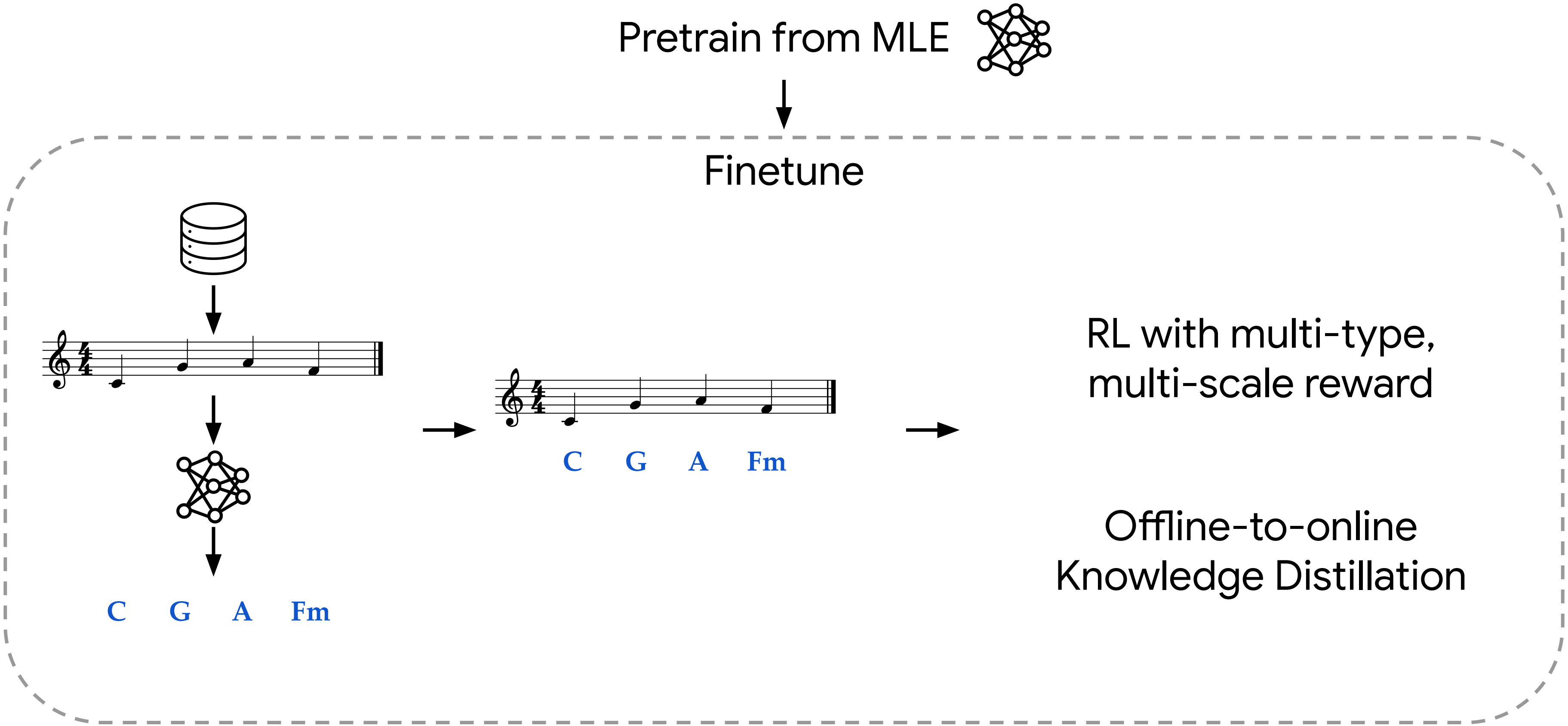
Finetune



Extra Objectives



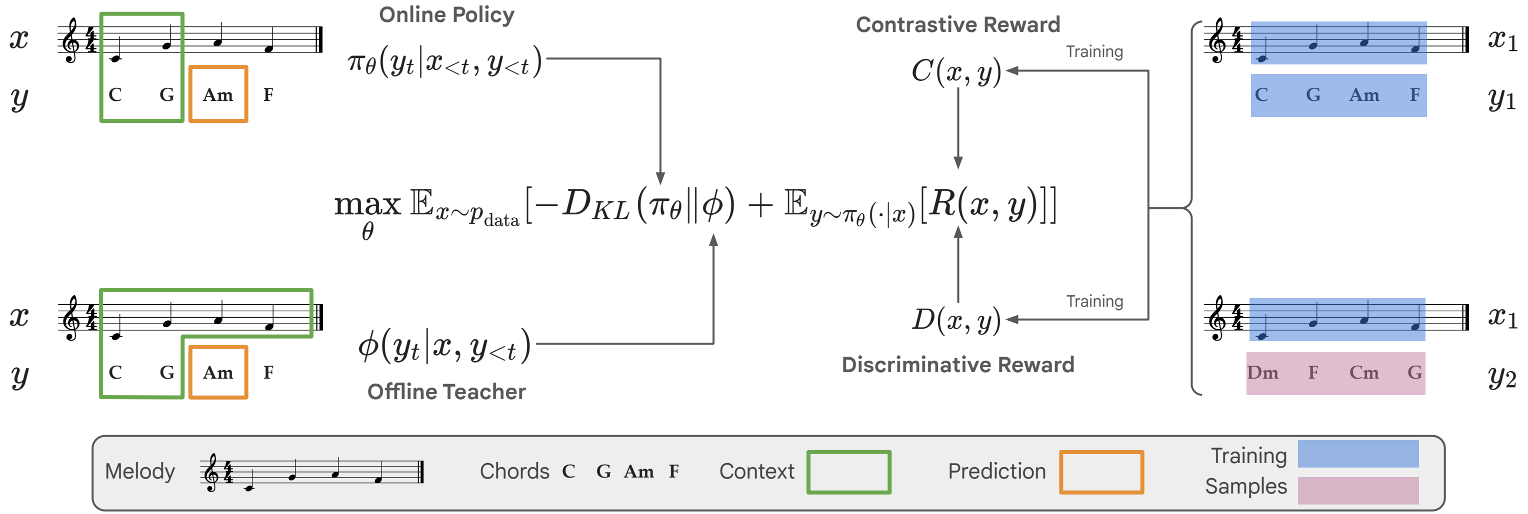
Extra Objectives



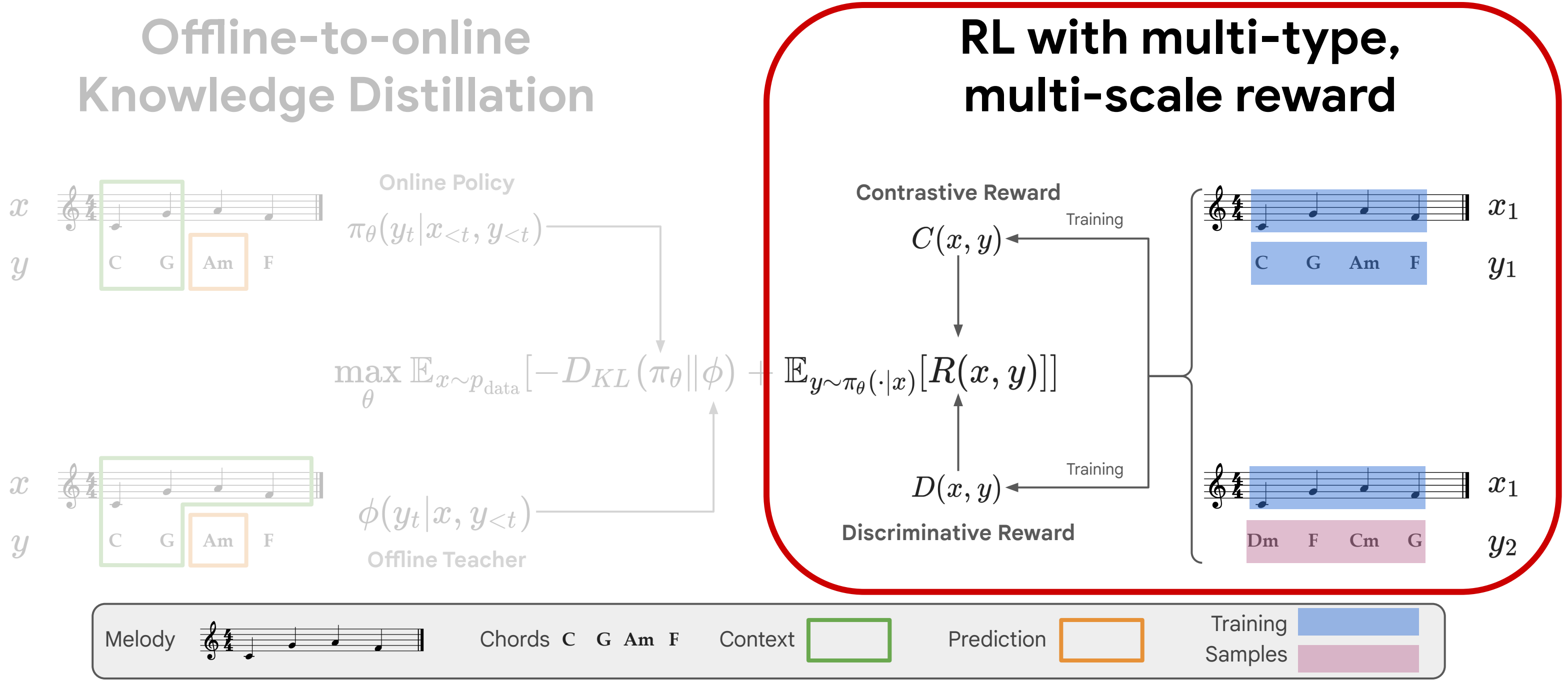
ReaLchords

Offline-to-online Knowledge Distillation

RL with multi-type, multi-scale reward

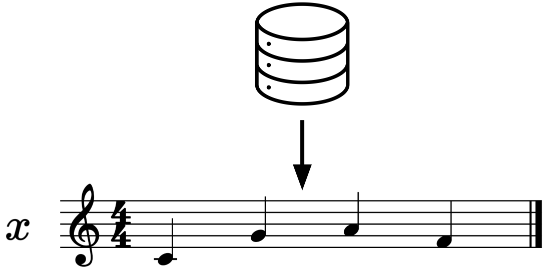


Reinforcement Learning

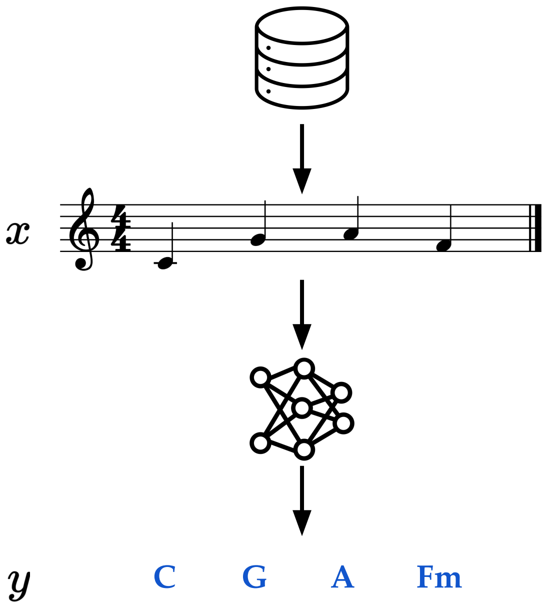


Reinforcement Learning

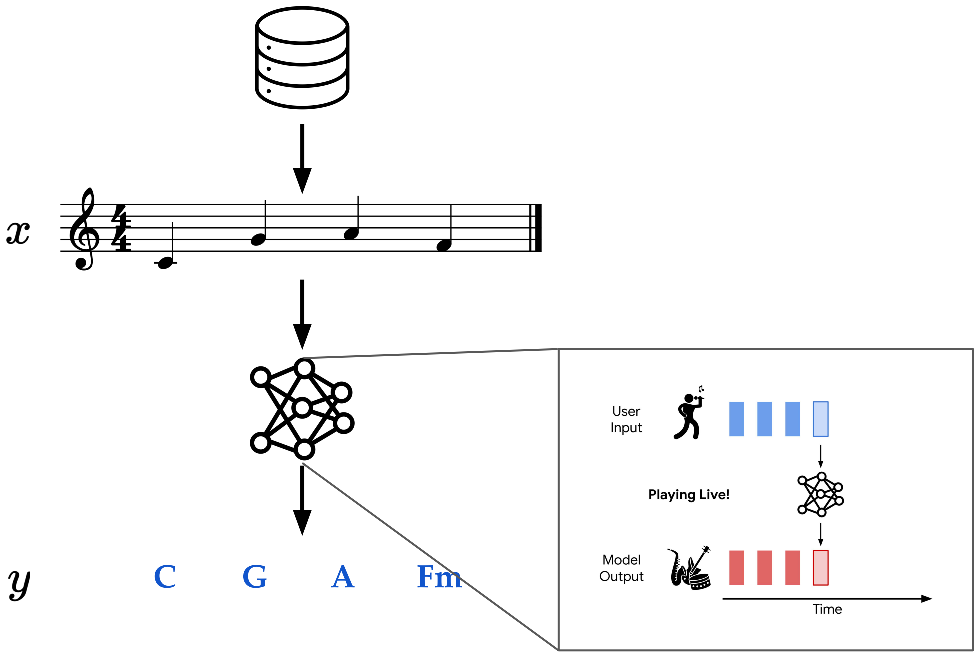
Reinforcement Learning



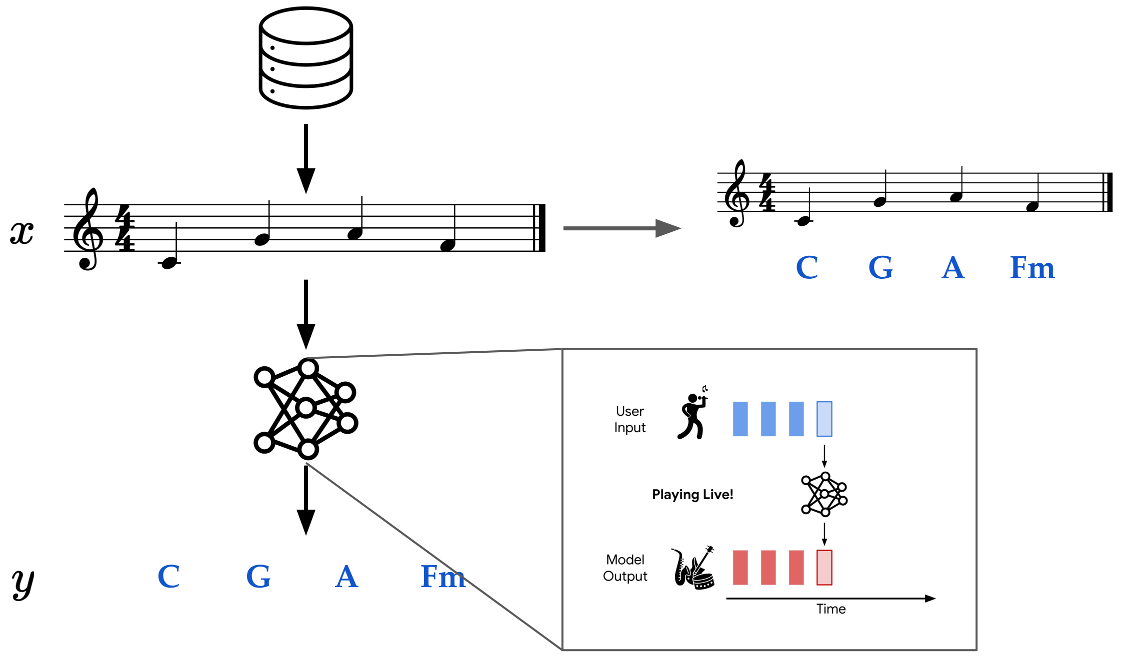
Reinforcement Learning



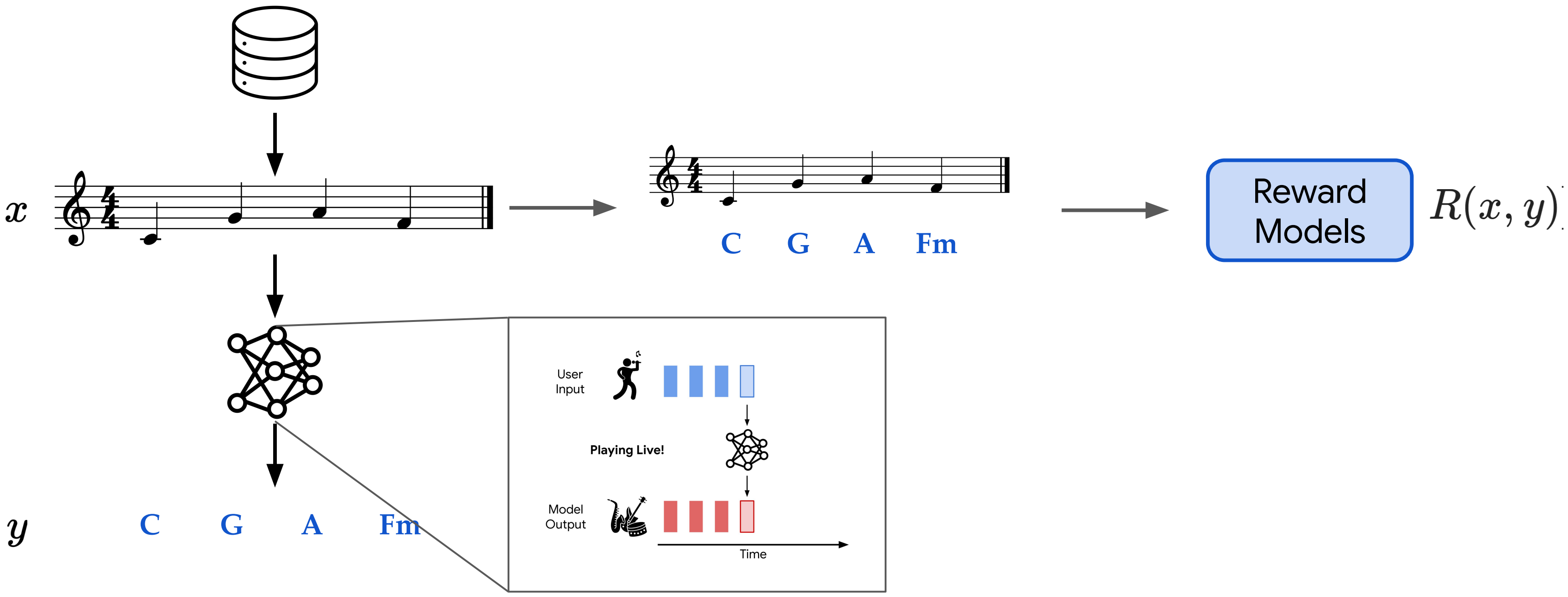
Reinforcement Learning



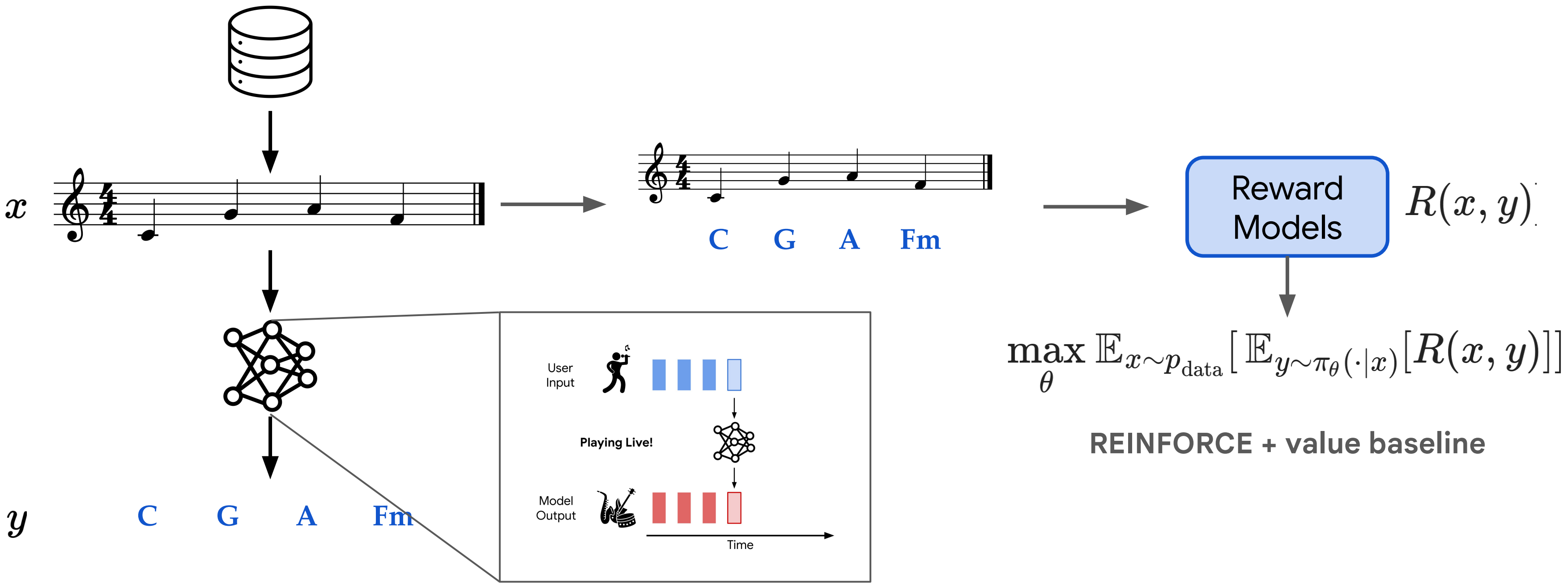
Reinforcement Learning



Reinforcement Learning



Reinforcement Learning



$$\max_{\theta} \mathbb{E}_{x \sim p_{\text{data}}} \left[\mathbb{E}_{y \sim \pi_\theta(\cdot|x)} [R(x, y)] \right]$$

REINFORCE + value baseline

Reward Model

- Dataset only contains melody & chords

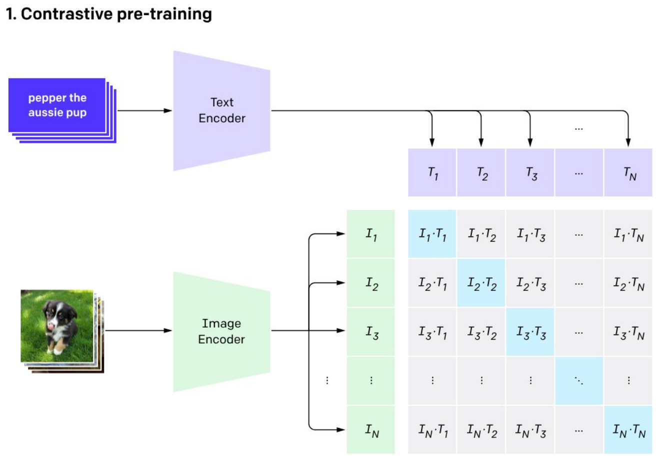
Reward Model

- Dataset only contains melody & chords
- Measure on how well model plays with input

Reward Model

- Dataset only contains melody & chords
- Measure on how well model plays with input

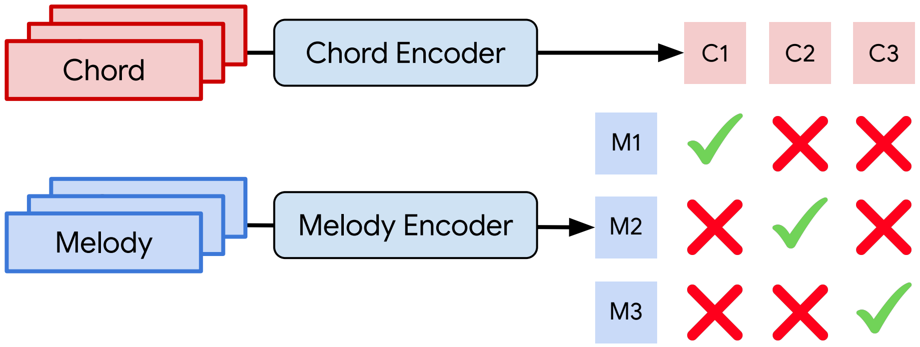
→ Train similarity measurement model via self-supervision



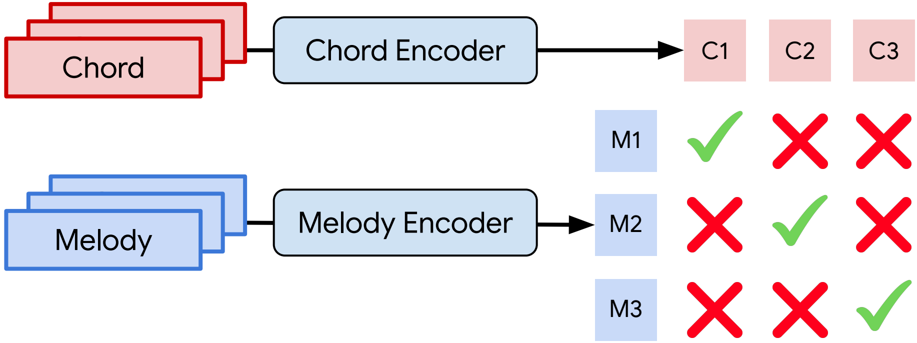
Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In International conference on machine learning (pp. 8748-8763). PMLR.

Contrastive Reward Model

- Train with contrastive objective
- Reward model takes in a complete episode of melody or chord
- Reward as similarity between melody (input) and chord (output)

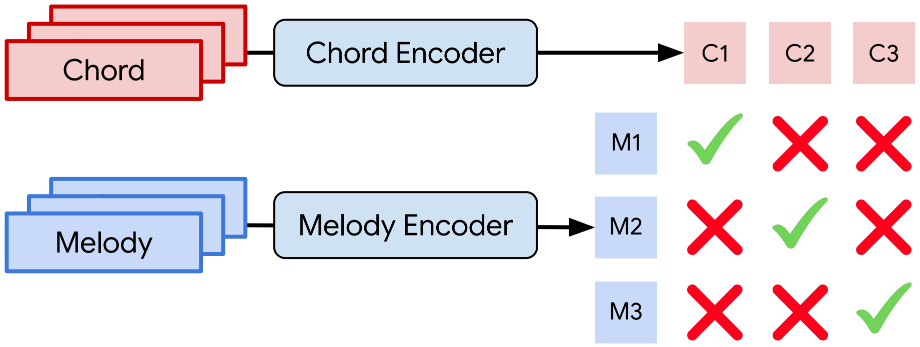


Contrastive Reward Model is **Not Enough**



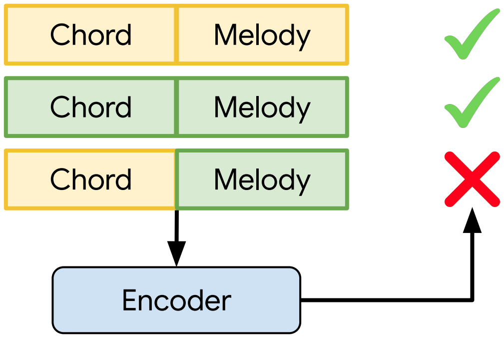
Contrastive Reward Model is **Not Enough**

- **Limited aspect:** Reward overemphasize overall harmonic
- **Granularity:** any single error could result in big drop in reward



Discriminative Model: reward with different perspective

- Train the self-supervise contrastive task as binary classification
- Reward as probability of classification
- Measure reward more on synchronization
- *Ensembling same model also boost performance

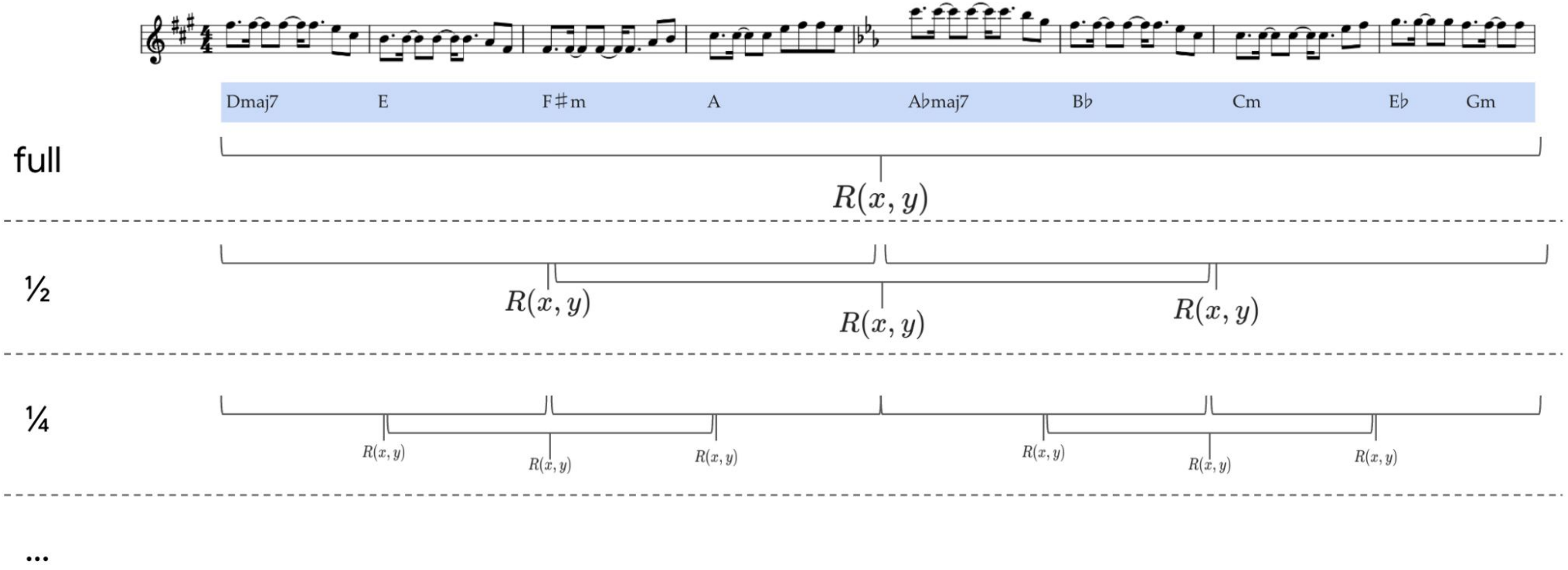


Multi-scale Reward for Better Credit Assignment

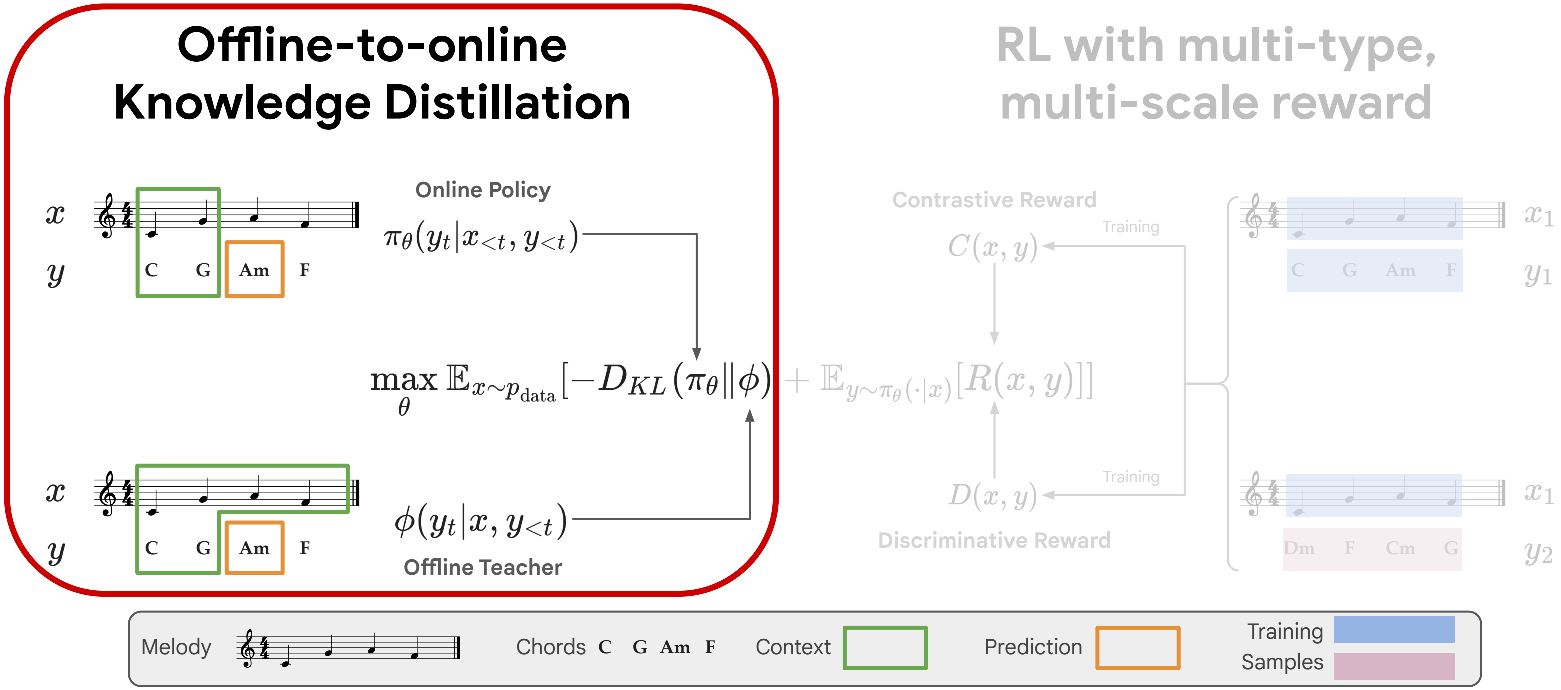
- Train reward models with smaller input context
- Ensemble multi-scale models

Multi-scale Reward for Better Credit Assignment

- Train reward models with smaller input context
- Ensemble multi-scale models
- $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}\}$ of the full context length, 50% overlap window

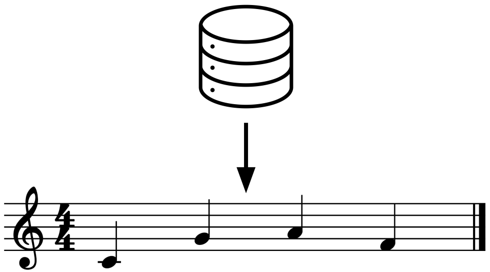


Knowledge Distillation

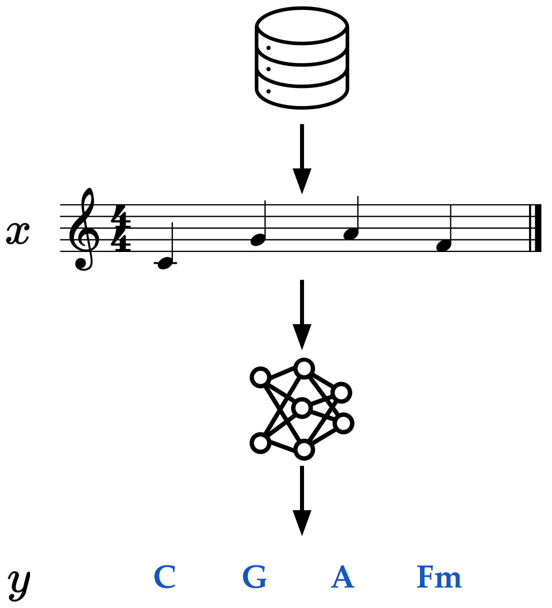


Knowledge Distillation

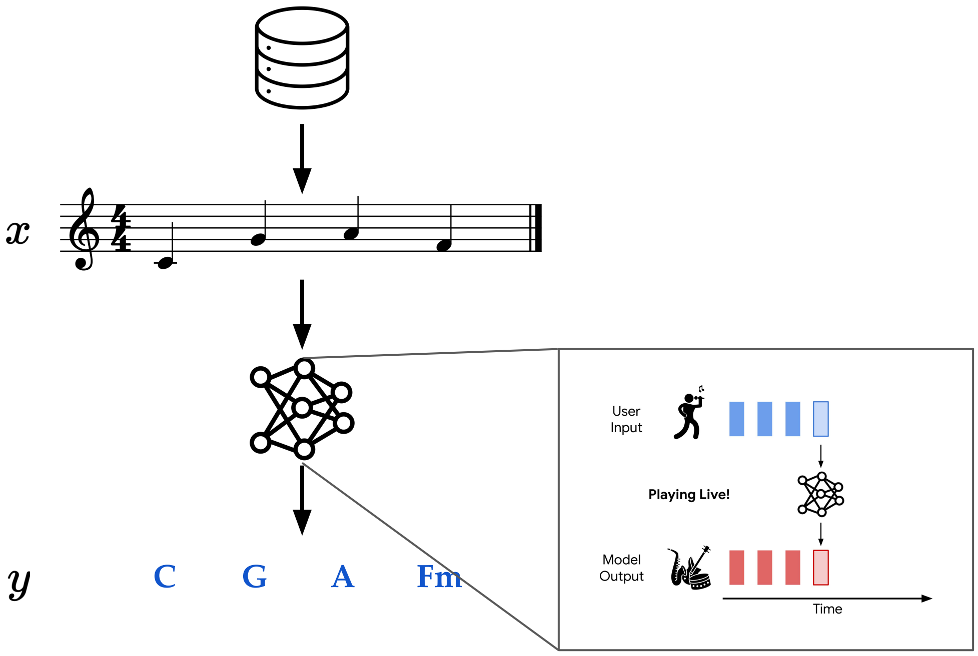
x



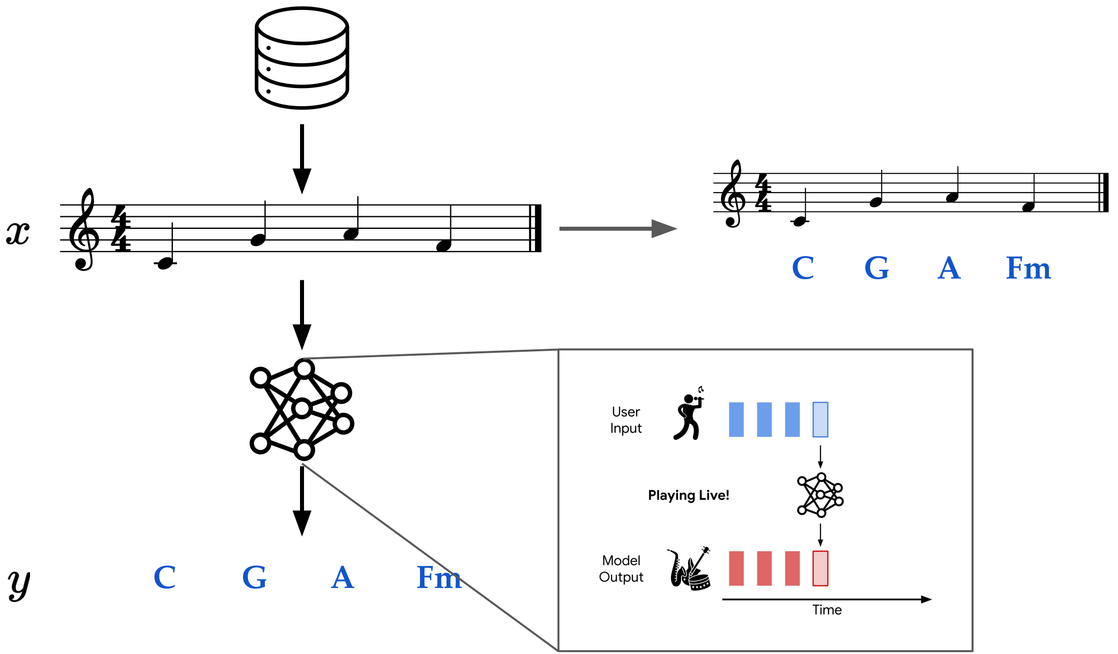
Knowledge Distillation



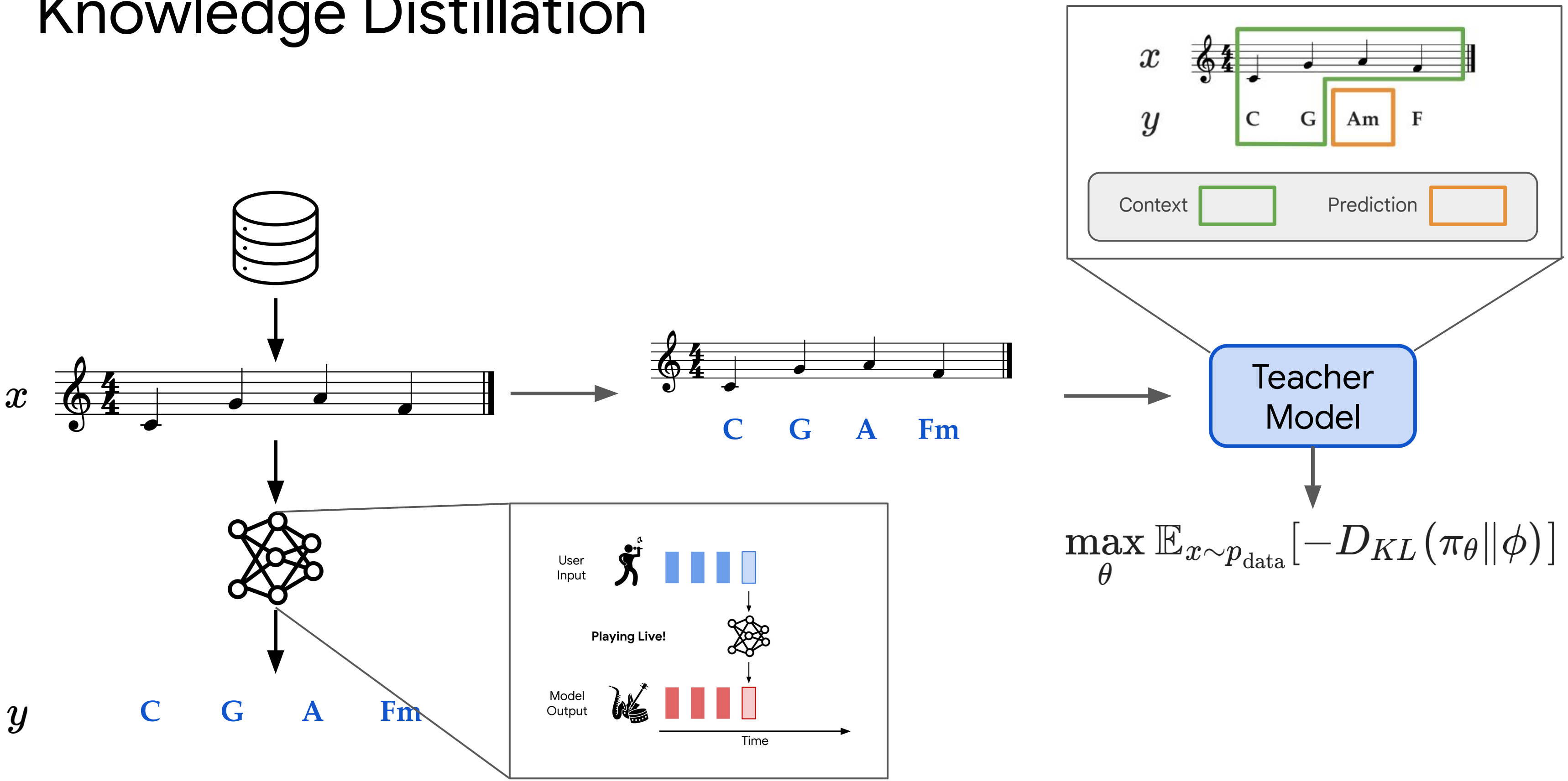
Knowledge Distillation



Knowledge Distillation

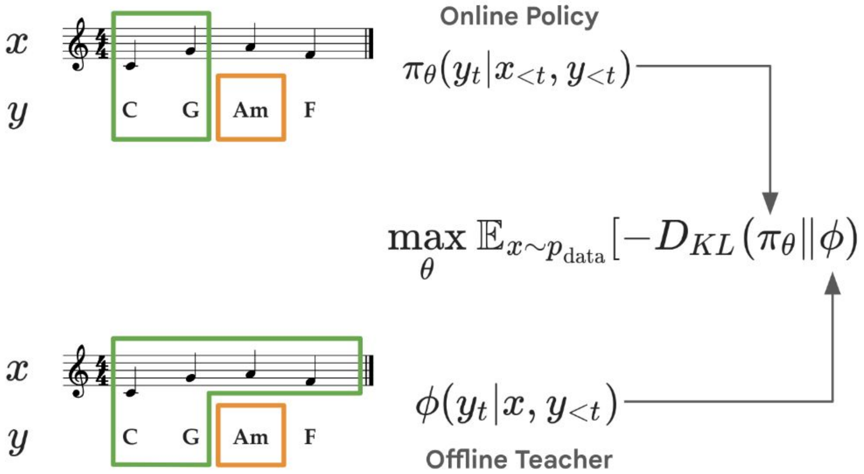


Knowledge Distillation

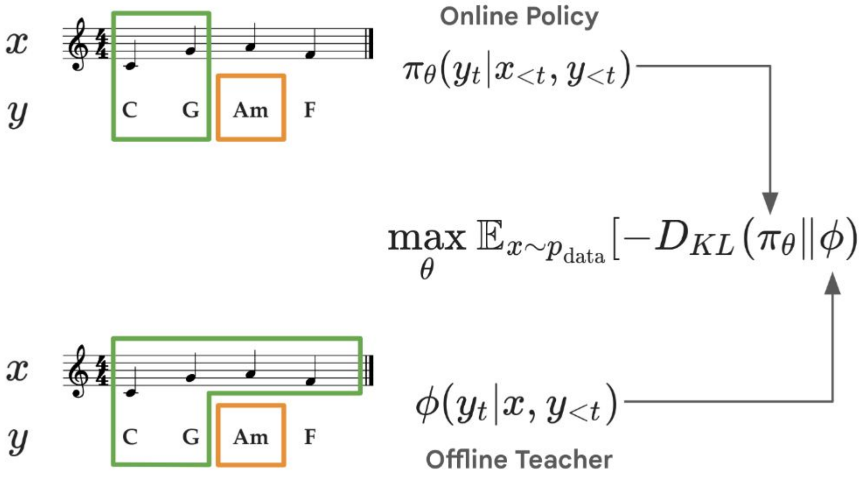


$$\max_{\theta} \mathbb{E}_{x \sim p_{\text{data}}} \left[-D_{KL}(\pi_\theta \| \phi) \right]$$

Offline-to-online distillation with previlliged information

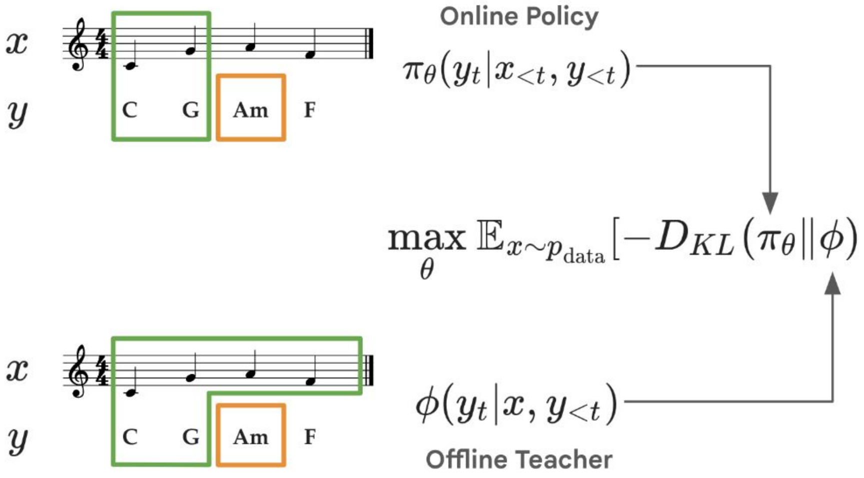


Offline-to-online distillation with previlliged information



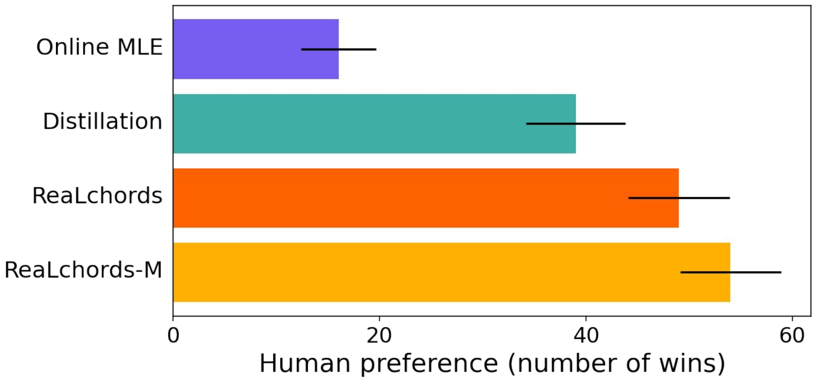
- Traditional knowledge distillation applies KL on **data samples** between **different model size**

Offline-to-online distillation with previlliged information

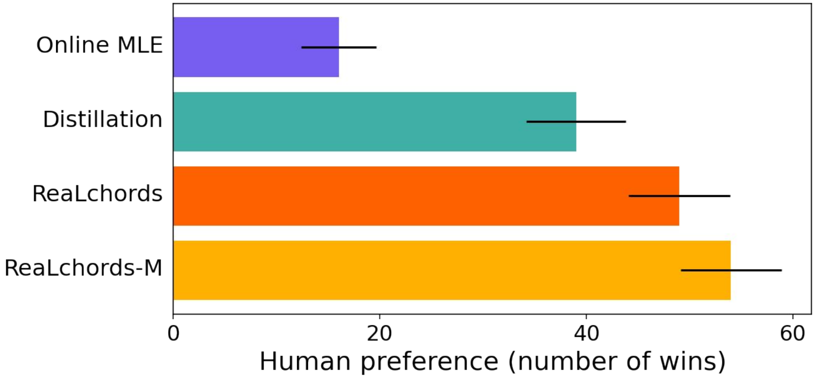


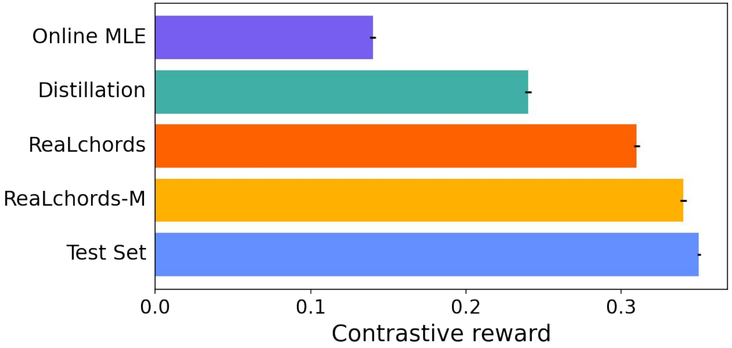
- Traditional knowledge distillation applies KL on **data samples** between **different model size**
- We apply KL on samples **generated from interaction** between **offline teacher and online policy**

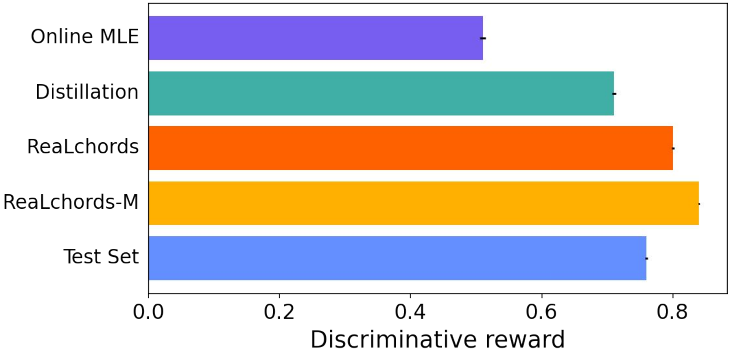
Better Generation Quality w/ RL & Knowledge Distillation



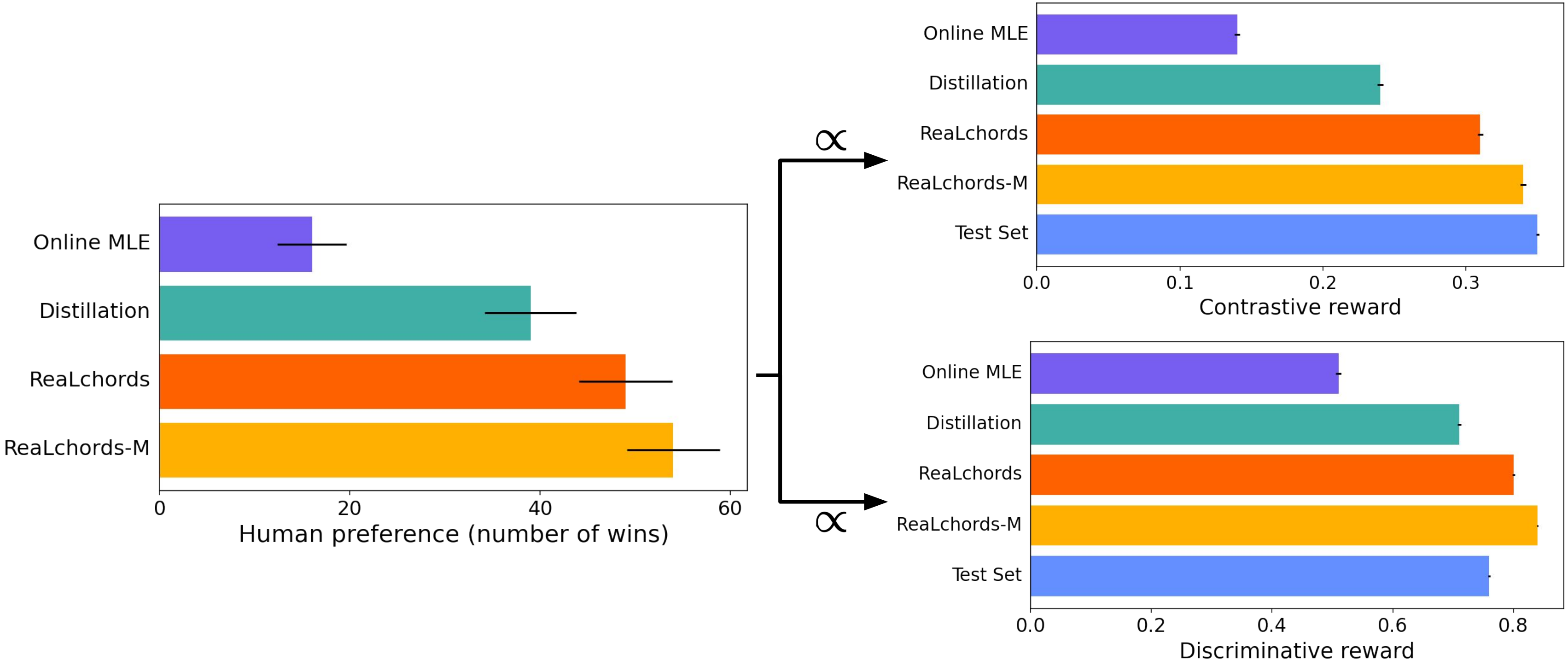
Better Generation Quality & Perceptually Aligned Reward





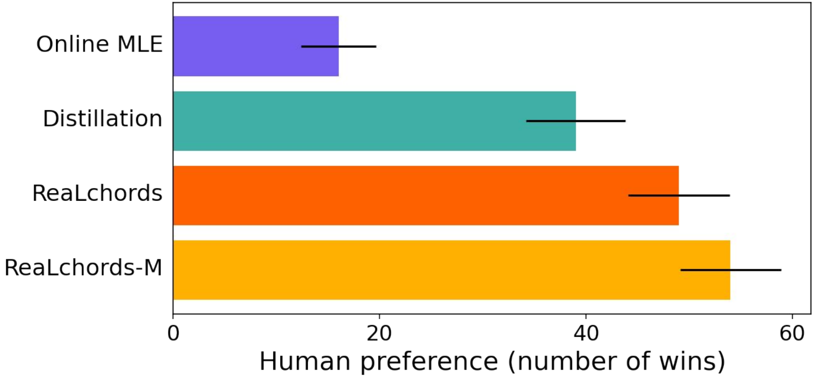


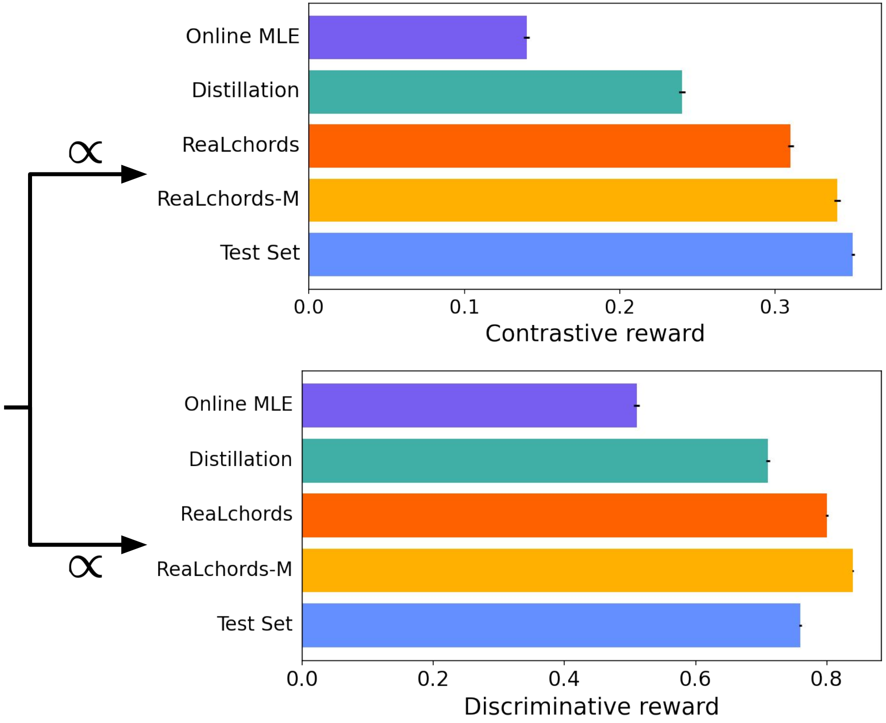
Better Generation Quality & Perceptually Aligned Reward



Better Generation Quality & Perceptually Aligned Reward

Human preference
$\propto$ **Contrastive reward**
$\propto$ **Discriminative reward**





Adaptive Accompaniment with ReaLchords

Check more audio samples here:
https://storage.googleapis.com/realchords/index.html