Real-Time Piano Accompaniment Using Kuramoto Model for Human-Like Synchronization

Kit Armstrong, Ji-Xuan Huang, Tzu-Ching Hung, Jing-Heng Huang & Yi-Wen Liu

Context

("Western classical music")

Composer – Performer – Listener

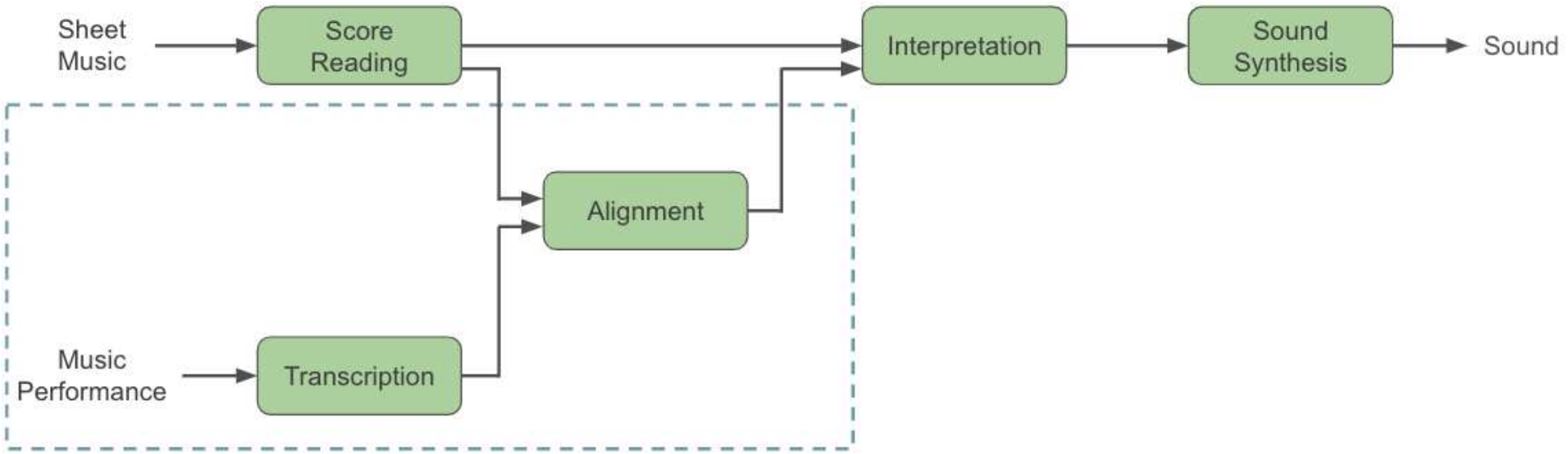
Context

(“Western classical music”)

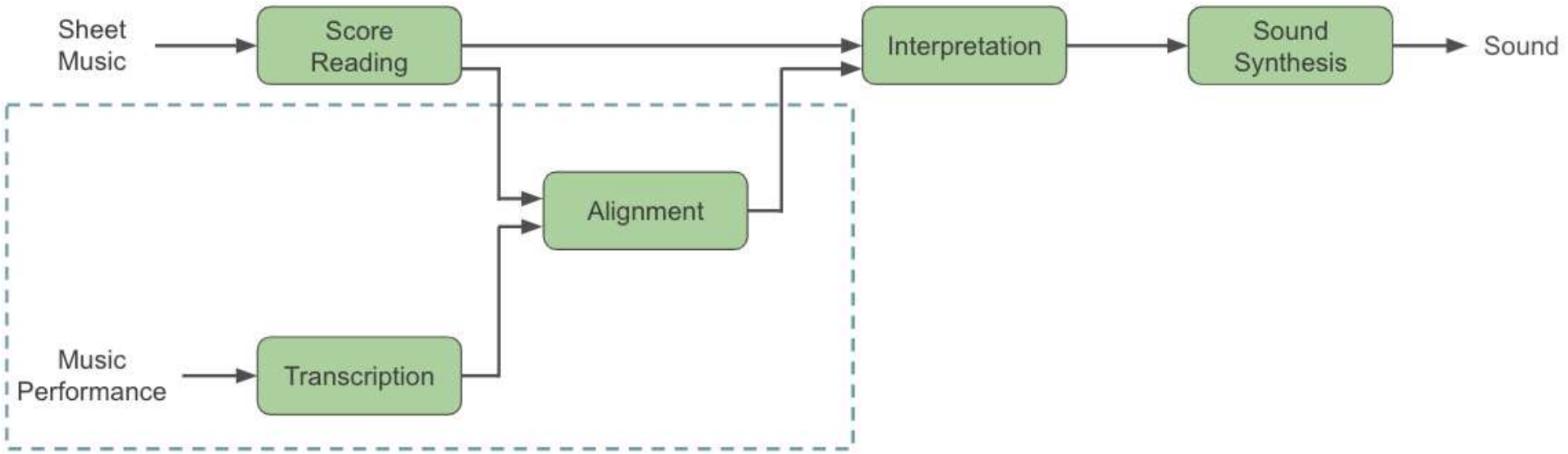
Composer – Performer – Listener

Create an AI performer

Context


Sheet
Music
Score
Reading
Interpretation
Sound
Synthesis
Sound
Alignment
Music
Performance
Transcription

Context



Score reading: sheet music → a machine-friendly representation
Transcription: musical performance → a machine-friendly representation
Score following: musical performance + score → score-aligned performance
Interpretation: score + score-aligned performance → MIDI or similar
Sound synthesis: interpretation → sound

Focus: Interpretation

Prevalence:

- Music-notation programs
  like Finale, Sibelius, MuseScore, etc.
- Useful tool for composers
- Not suitable for performance

Focus: Interpretation

Prevalence:

- Music-notation programs
  like Finale, Sibelius, MuseScore, etc.
- Useful tool for composers
- Not suitable for performance

Levels of advancement:

- Basic, in essence MusicXML to MIDI
- Algorithmic expressiveness
- Ongoing attempts with machine learning

Focus: Interpretation

Challenges:

- Deformation for natural performance

Focus: Interpretation

Challenges:

- Deformation for natural performance
- Understanding score indications

Focus: Interpretation

Challenges:

- Deformation for natural performance
- Understanding score indications
- Synchronization in an ensemble

Model

Scope:

- Digital piano
- 1 person (“input”) + AI “accompaniment”
- Precisely defined score

Goal:

- Human-like time synchronization

Model

Scope:

- Digital piano
- 1 person (“input”) + AI “accompaniment”
- Precisely defined score

Goal:

- Human-like time synchronization

Collaboration more than synchronization

Model

Human-like time synchronization

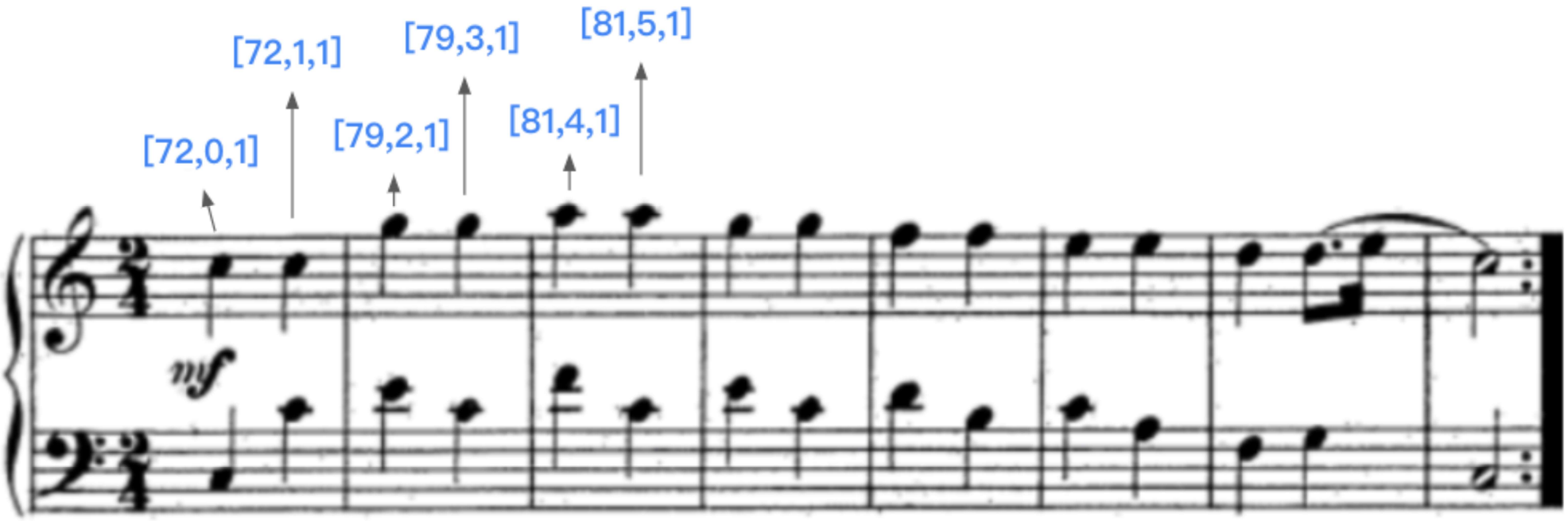
Intuition:

“I'm early” → Slow down
“I'm late” → Speed up

Model

Human-like time synchronization

Introduce the concept of **score position**



Model

Human-like time synchronization

A performance is a function: time → score position

$f(t)$ = 2π * (beat number reached at time t)

Model

Human-like time synchronization

A performance is a function: time → score position

f(t) = 2π * (beat number reached at time *t*)

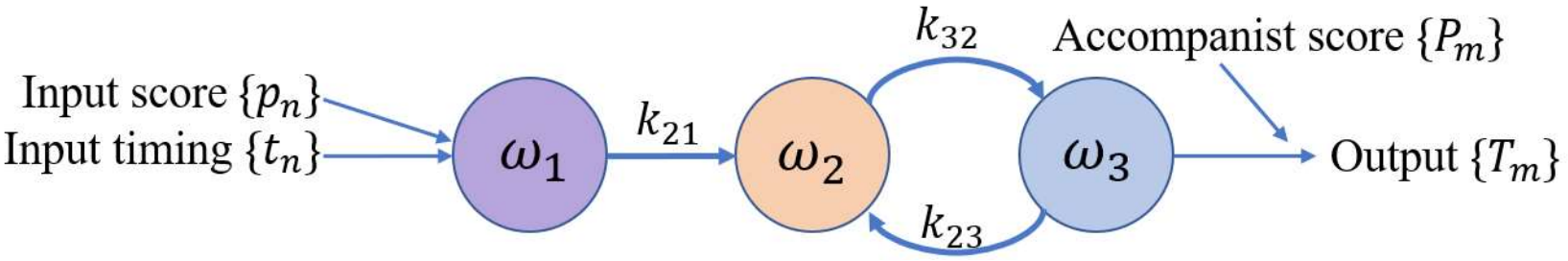
{input timings}, {input score positions} determine *f*

Input is one such function,
accompaniment is another function

Model

Intuition:

“I'm early” → Slow down
“I'm late” → Speed up



$$\frac{d\theta_i(t)}{dt} = \sum_{j \neq i} k_{ij} \sin\left(\theta_j(t) - \theta_i(t)\right) + \Omega_i(t)$$

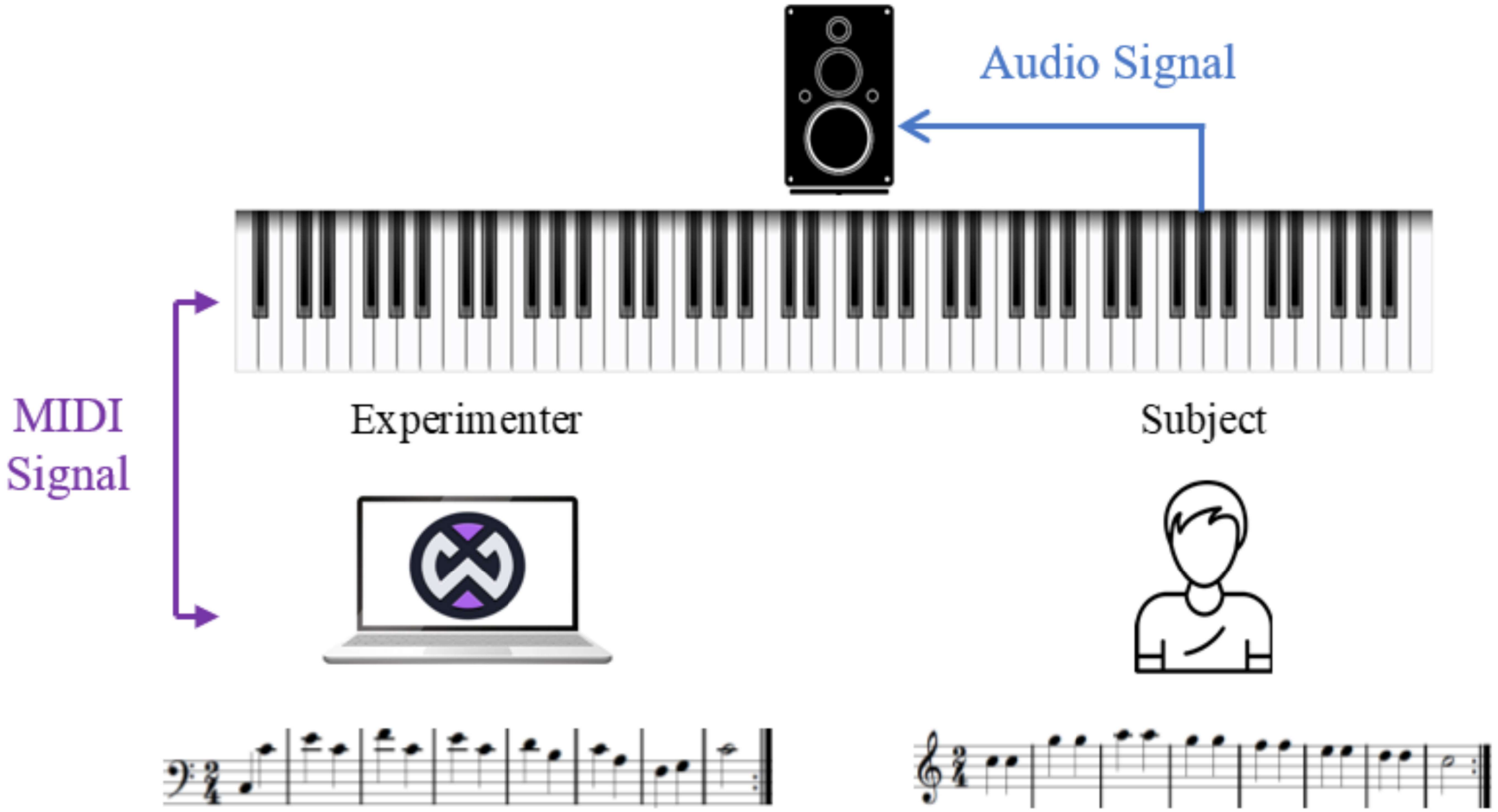
Model

Predictive nature:

- extrapolate ω_3 to predict next outputs
- additional learned parameter: “reaction time”
- each new input rewrites all future predictions

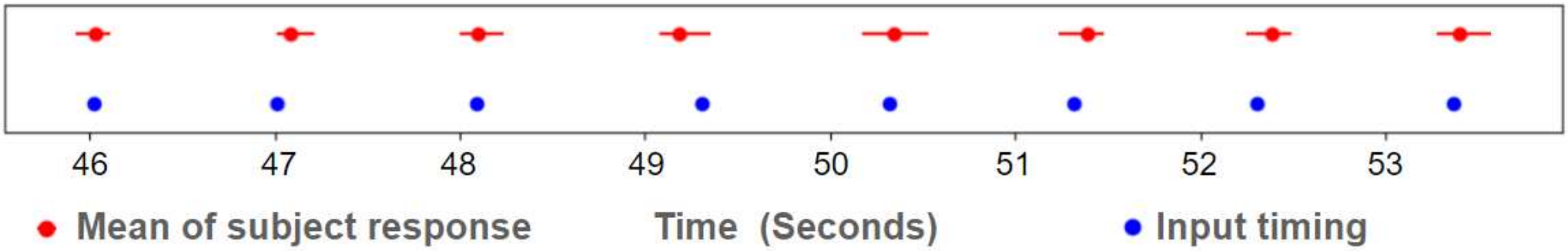
Model training

Capturing human behavior



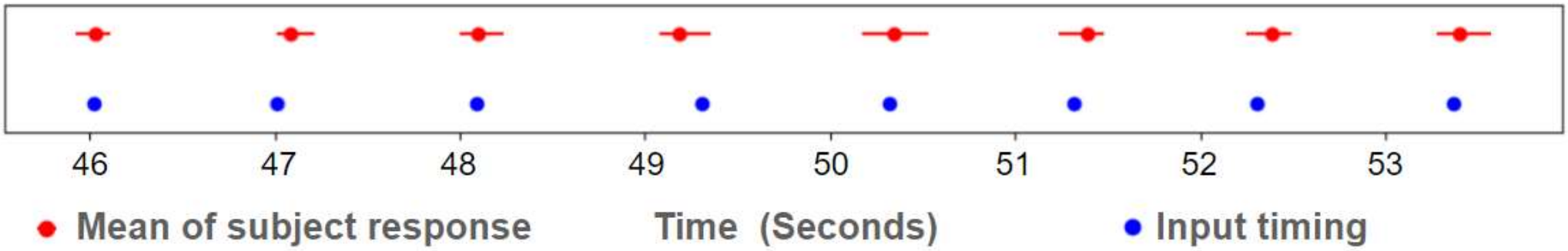
Model training

Capturing human behavior



Model training

Capturing human behavior



Choose parameters such that
the model performs most similarly

Architecture

- input reading thread

  writes all MIDI messages to *Input Queue*

- calculation thread

  listens to *Input Queue*,
  determines outputs of form (note, velocity, time),
  and writes them to *Output Queue*

- output thread

  listens to *Output Queue*,
  sends MIDI messages at corresponding time

Demonstration

Additional elements:

- Simple velocity matching (running average)
- Basic error-correction

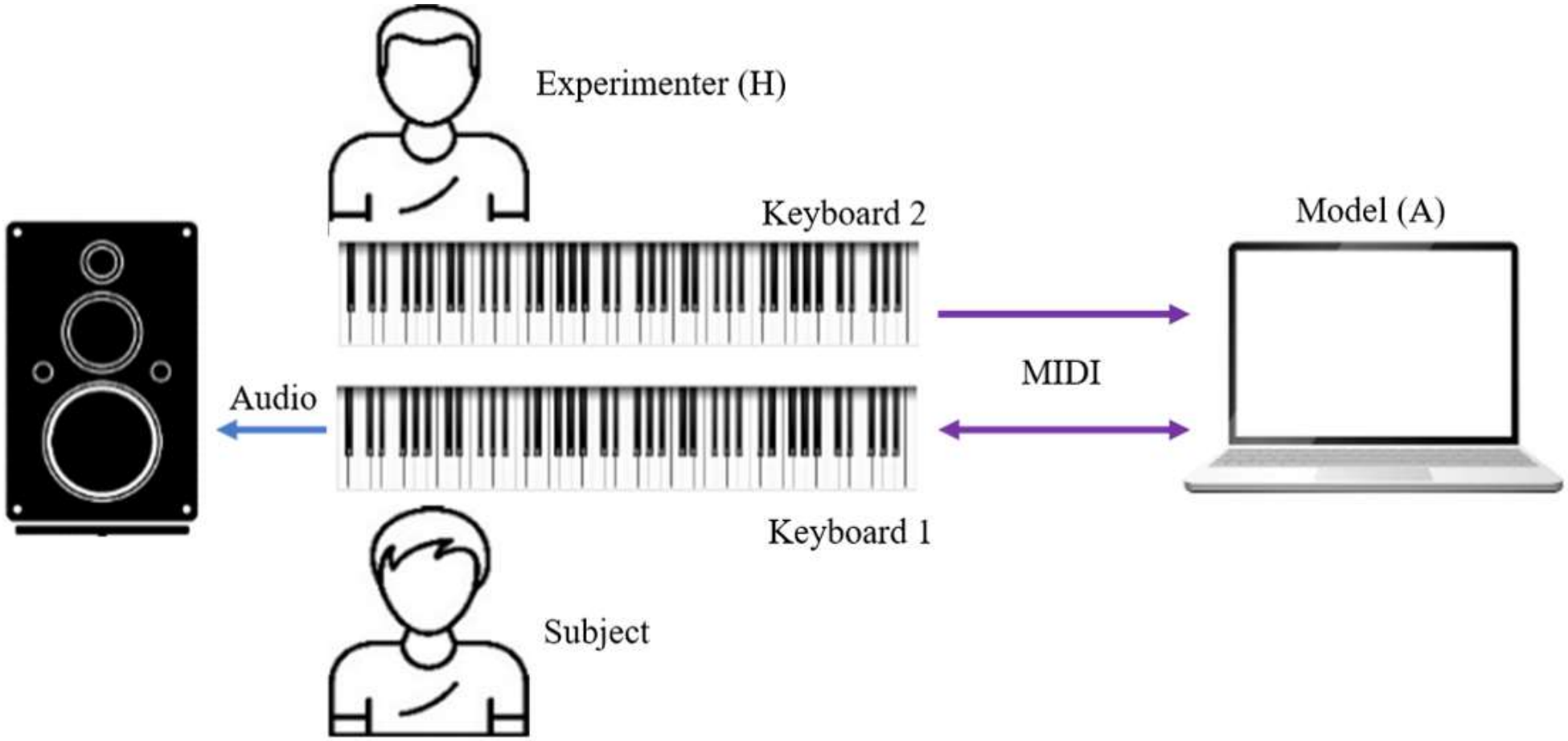
Testing

“Turing test”

Can people tell the difference between Human (H) and AI (A)?

Environment:
Create identical setup for H and A

Testing


Experimenter (H)
Keyboard 2
Model (A)
MIDI
Audio
Keyboard 1
Subject

Testing

- **1.**
  W. A. Mozart: *"Twinkle, Twinkle, Little Star"* K. 265, Theme
- **2.**
  W. A. Mozart: *"Twinkle, Twinkle, Little Star"* K. 265, Variation II
- **3.**
  J. S. Bach/C. Gounod: *"Ave Maria, Méditation sur le Prélude de Bach"*

Testing

Results:

12 participants, 80 trials

Total: 59% correct guesses

Problems & Limitations

Problem: Output delay

Getting the OS to reliably send a MIDI message at a pre-determined future time

- Misrepresents the model
- Bumpy effect
- Cascading slow-down
- Worse when the output has many notes

Problems & Limitations

Limitations:

- Only rhythm is taken into account

Problems & Limitations

Limitations:

- Only rhythm is taken into account

- Each note carries the same weight or salience

Problems & Limitations

Limitations:

- Only rhythm is taken into account

- Each note carries the same weight or salience

General limitations:

- No intrinsic musicality

- Primitive score reading

Taking stock

- This model is conceptually simple,
  and its scope is very limited.
- Playing with it is surprisingly satisfying,
  especially compared to “perfect accompaniment”.
- New trial suggests this may not be thanks to the Kuramoto model, but rather the predictive nature.
- We are devising a new model with larger scope and different internal workings.