

# Music Understanding and the Future of Music Performance

Roger B. Dannenberg

Professor of Computer Science, Art, and Music  
Carnegie Mellon University

## Why Computers and Music?

- Music in every human society!
- Computing can make music:
  - More Fun
  - More Available
  - Higher Quality
  - More Personal



© 2013 Roger B. Dannenberg

2

## My Background

- Always interested in math and music and making things
- Discovered synthesizers in high school
- Discovered computers about the same time
- Discovered computer music in college
- Research motivated by musical experience:
  - Computer accompaniment
  - Expressive programming languages for music
  - Audacity
  - ... current work

## Overview

- Introduction
- 
- How Is Computation Used in Music Today?
  - New Capabilities:
    - What Can Computers Do Tomorrow?
  - What Will Music Be Like in the Future?

Carnegie Mellon

## How Is Computation Used in Music Today?



Indabamusic.com



making-music.blogspot.com



<http://venturebeat.com/>



© 2013 Roger B. Dannenberg


5

Carnegie Mellon

## Music Computation Today

Production: digital recording, editing, mixing

- Nearly all music production today...
  - Records audio to (digital) disk
  - Edit/manipulate audio digitally
    - Equalization
    - Reverberation
  - Convert to media:
    - CD
    - MP3
    - Etc.



protools.com


© 2013 Roger B. Dannenberg

6


Carnegie Mellon

# Music Computation Today

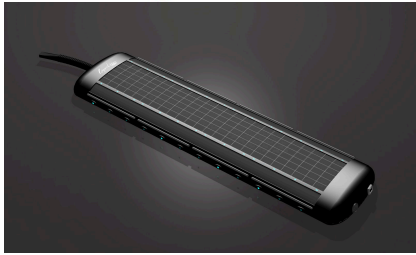
## Musical Instruments: synthesizers and controllers



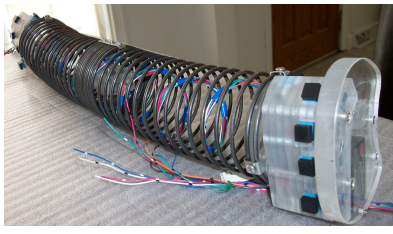
Synthesizer (Solaris)



Drum Machine (Yamaha)



Linnstrument (Roger Linn)



Sonic Spring (Tomas Henriques)


© 2013 Roger B. Dannenberg

7


Carnegie Mellon

# Music Computation Today

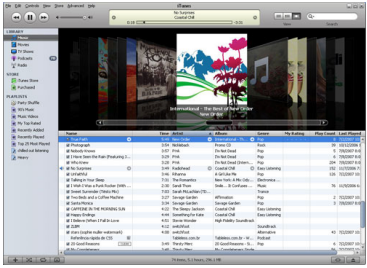
## Distribution: compression, storage, networks



Napster




Apple iPod



Apple iTunes

Now with Unlimited Space for Music  
Keep Your Music Library Online and Enjoy It Anywhere with Cloud Player



amazon cloud drive    amazon cloud player

Amazon Cloud Player


© 2013 Roger B. Dannenberg

8

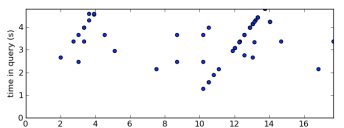
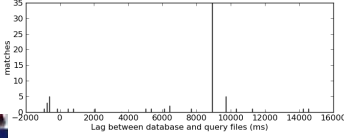
Carnegie Mellon

# Music Computation Today

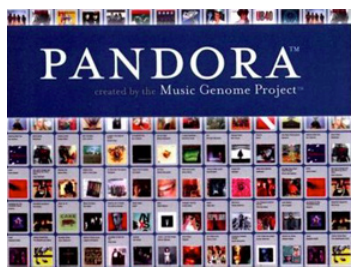
## Search, recommendation, music fingerprinting



Google Music China

Music Fingerprinting



Pandora  
Music Recommendation

© 2013 Roger B. Dannenberg

9

Carnegie Mellon

# Overview

- Computer Music Introduction
- How Is Computation Used in Music Today?

---

- New Capabilities:  
What Can Computers Do Tomorrow?
- What Will Music Be Like in the Future?

© 2013 Roger B. Dannenberg

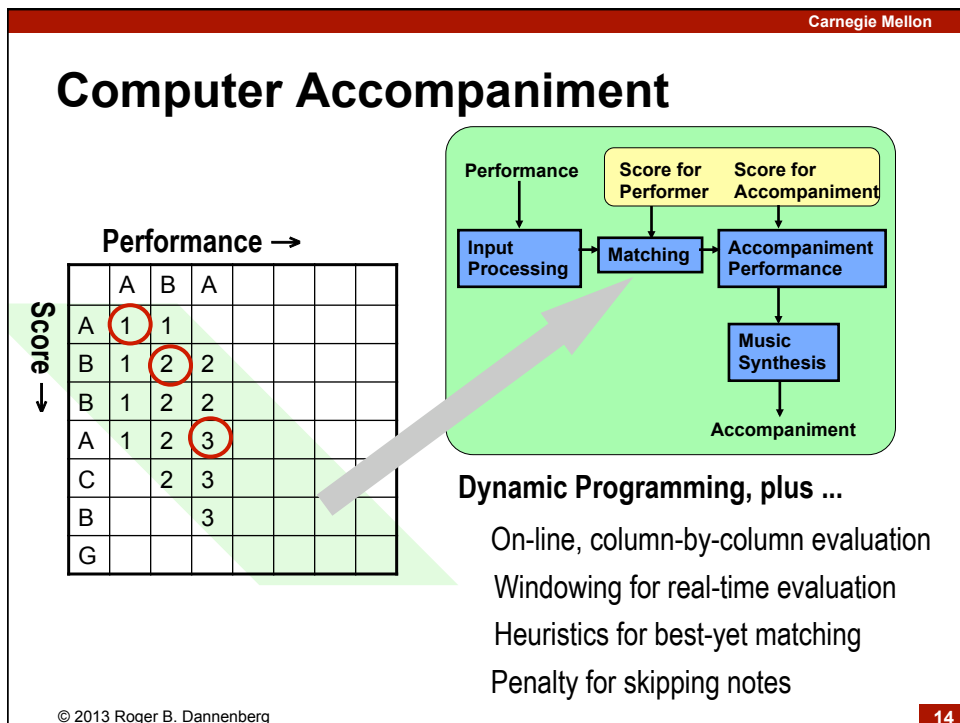
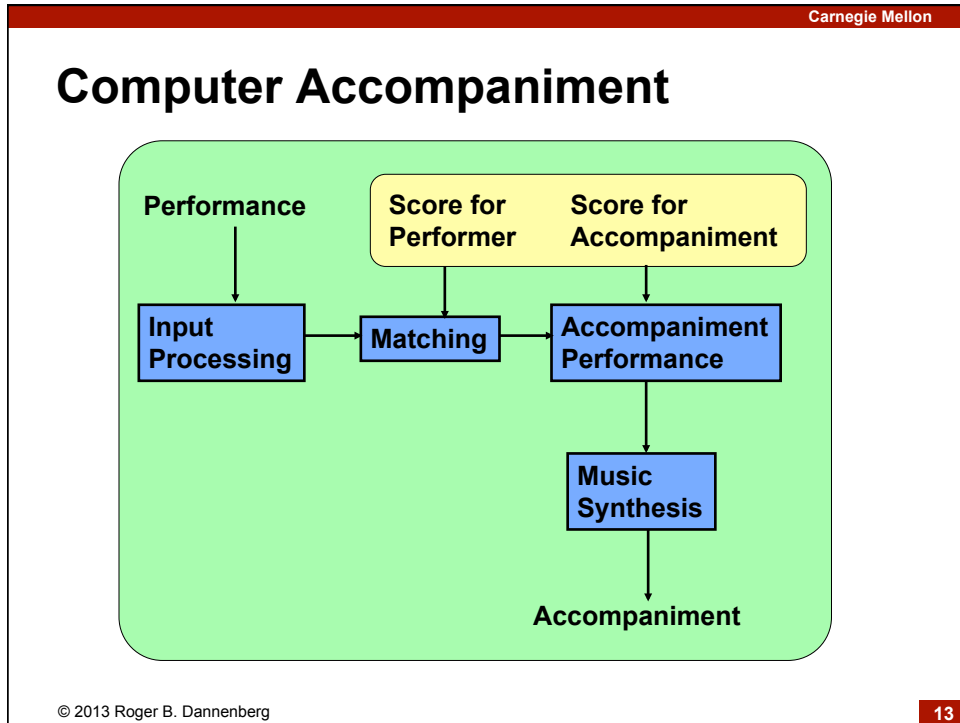
10

## New Capabilities: What Can Computers Do Tomorrow?

- Computer accompaniment
- Style classification
- Score alignment
- Onset detection
- Sound synthesis

## Accompaniment Video





Carnegie Mellon

## Computer Accompaniment

**Rule-based system:**

E.g. If matcher is confident and accompaniment is ahead < 0.1s, stop until synchronized.

If matcher is confident and accompaniment is behind < 0.5s, speed up until synchronized.

```

graph TD
    Performance --> IP[Input Processing]
    IP --> Matching
    Matching --> AP[Accompaniment Performance]
    AP --> MS[Music Synthesis]
    MS --> Accompaniment
    subgraph Scores [ ]
        S1[Score for Performer]
        S2[Score for Accompaniment]
    end
    S1 --> Matching
    S2 --> AP
    
```

The diagram shows a process flow for computer accompaniment. It starts with 'Performance' input to 'Input Processing', which leads to 'Matching'. 'Matching' also receives 'Score for Performer' and outputs to 'Accompaniment Performance'. 'Accompaniment Performance' also receives 'Score for Accompaniment' and outputs to 'Music Synthesis', which finally produces the 'Accompaniment'.

15

© 2013 Roger B. Dannenberg

Carnegie Mellon

## Vocal Accompaniment

- Lorin Grubb's Ph.D. (CMU CSD)
- Machine learning used to:
  - Learns what kinds of tempo variation are likely
  - Characterize sensors
    - When is a notated G sensed as a G#?
- Machine learning necessary for good performance

A small video inset showing a woman with dark hair singing into a microphone.

16

© 2013 Roger B. Dannenberg



Carnegie Mellon

## Vocal Accompaniment



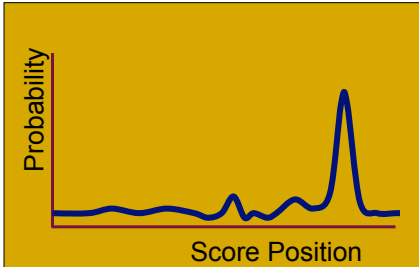
© 2013 Roger B. Dannenberg

17

Carnegie Mellon

## How It Works

- Score position modeled as a probability density function
- Bayesian update rule:  
 $P(s|o) \propto P(o|s)P(s)$
- $P(o|s)$  is e.g. "probability of observing pitch G if the score says play an A."  
 Simple statistics on labeled training data.
- Prior  $P(s)$  by fast *convolution* with a log normal (describes tempo and tempo variation)



© 2013 Roger B. Dannenberg

18

Carnegie Mellon

## Commercial Implementation



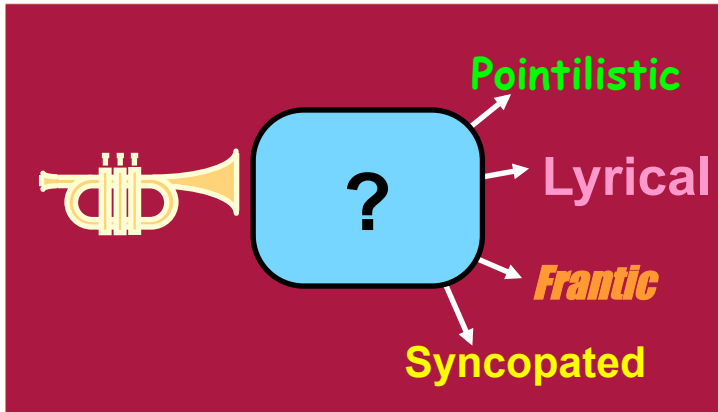
[https://ot.partner-streaming.com/makemusic/wm\\_03\\_L.pov](https://ot.partner-streaming.com/makemusic/wm_03_L.pov)    [https://ot.partner-streaming.com/makemusic/wm\\_04\\_L.pov](https://ot.partner-streaming.com/makemusic/wm_04_L.pov)

© 2013 Roger B. Dannenberg

19

Carnegie Mellon

## Style Classification: Listening to Jazz Styles



Pointilistic

Lyrical

Frantic

Syncopated

© 2013 Roger B. Dannenberg

20

Carnegie Mellon

## Jazz Style Recognition



© 2013 Roger B. Dannenberg

21

Carnegie Mellon

## Techniques

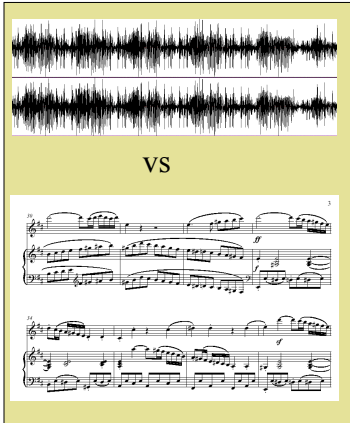
- Extract features from audio:
  - Note density
  - Mean & Std. Dev. of pitch range
  - Mean & Std. Dev. of pitch intervals
  - Silence vs. Sounding ("duty factor")
  - ... and many more
- Features over 5-second windows
- Standard Classifiers (Naive Bayes, Linear, Neural Net)

© 2013 Roger B. Dannenberg


22

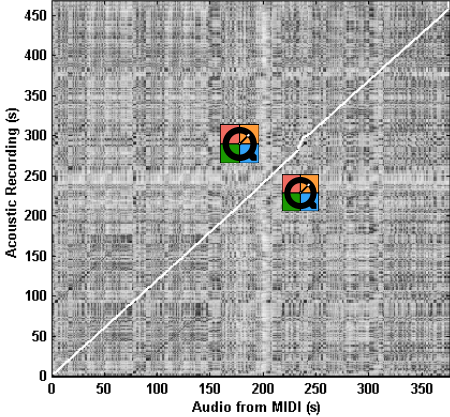
Carnegie Mellon

## Polyphonic Audio-to-Score Alignment



VS

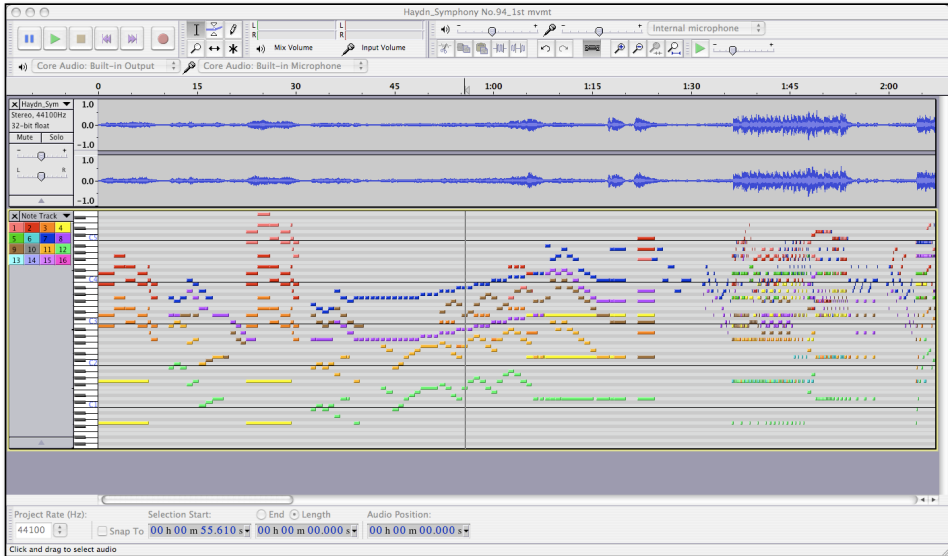




© 2013 Roger B. Dannenberg23

Carnegie Mellon

## Audacity Editor with Automatic Audio-to-MIDI Alignment

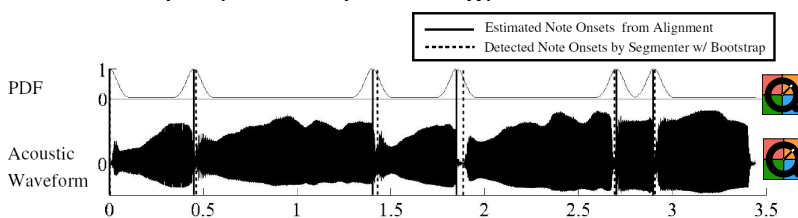


Project Rate (Hz): 44100  
Selection Start: 00 h 00 m 55.610 s  
End: 00 h 00 m 00.000 s  
Length: 00 h 00 m 00.000 s  
Audio Position: 00 h 00 m 00.000 s

## Finding Note Onsets

(How to segment music audio into notes.)

- Not all attacks are clean
- Slurs do not have obvious (or fast) transitions
- We can use score alignment to get a rough idea of where the notes are (~1/10 second)
- Then, machine learning can create programs that do an even better job (bootstrap learning).

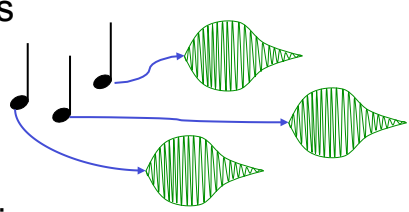


## Expressive Performance

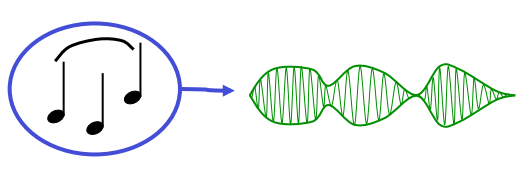
Carnegie Mellon

## Phrase-based Synthesis

Note-by-Note Synthesis



Phrase-based Synthesis

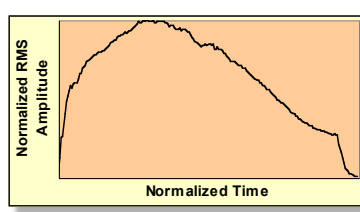


© 2013 Roger B. Dannenberg 27

The diagram illustrates two synthesis methods. The top part, 'Note-by-Note Synthesis', shows three individual notes on a staff. Blue arrows point from each note to a separate, individual waveform envelope. The bottom part, 'Phrase-based Synthesis', shows three notes on a staff grouped together under a blue slur. A blue arrow points from this group to a single, continuous waveform envelope that spans the duration of the entire phrase.

Carnegie Mellon

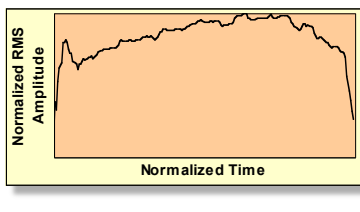
## Example Envelopes



Normalized RMS Amplitude

Normalized Time

Tongued Note



Normalized RMS Amplitude






Normalized Time

Slurred Note

© 2013 Roger B. Dannenberg 28

The diagram shows two amplitude envelopes. The top graph, labeled 'Tongued Note', shows a curve that rises to a peak and then decays. The bottom graph, labeled 'Slurred Note', shows a curve that rises to a sustained, relatively flat plateau and then decays.

## Synthesis Examples

- Good trumpet sounds, mechanically performed: 
- Same sounds, but performed with AI-based model of trumpet performance: 
- Another example: 
- Trumpet example from Ning Hu's thesis: 
- Bassoon example from Ning Hu's thesis: 

## Overview

- Computer Music Introduction
  - How Is Computation Used in Music Today?
  - New Capabilities:  
What Can Computers Do Tomorrow?
- 
- What Will Music Be Like in the Future?

## Human Computer Music Performance

- The most advanced computer music research is applied to esoteric art music.
  - There is a widespread practice of interactive computer (art) music
  - ... but relatively little sophistication in popular music
- OPPORTUNITY
  - State-of-the-art computer music systems for popular music performance
  - *Autonomous Intelligent Machine Musicians*

## Example

- Suppose you want to get together and play music ... BUT, you're missing a bass player.



credit: Green Day



## What Research Is Needed?

- Synchronization
  - Signal processing
  - Machine learning
  - Human interface
- Sketchy notation
  - Representation issues
- Improvisation
  - Models of style
- Sound Production
  - Phrase-based synthesis?
- Modularity/Systems issues
  - Real-time systems
  - Software architecture
- Interaction
  - HCI

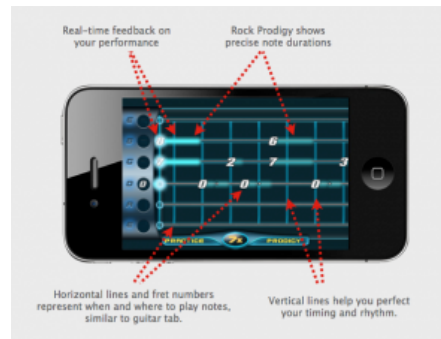


## Is There a Market? What's the Impact?

- \$8B annual US music sales
  - **Excluding** recordings, education, performances
- 5 million musical instruments per year
- Performance revenue is on the order of \$10B
- Recording revenue is similar; order of \$10B
- Approximately 1/2 of all US households have a practicing musician
- ... so *very roughly* \$10+B and 100M people!

## Rock Prodigy

- **Guitar Hero for Real Guitars**
- **Game design, content, animation, etc. by others**
- **(Play [Video](#))**
- **Unsolicited comment: "The best part about it is polyphonic pitch detection"**



## An Example



Carnegie Mellon

## Online, collaborative development of creative content is already here...

**What's New**  
 Music by **Bob Haggart**  
 Words by **Johnny Burke**  
 Performed by **Frank Sinatra, Helen Forrest**

© 2013 Roger B. Dannenberg

**37**

Carnegie Mellon

## What Will People Do With HCMP?

- Practice with virtual bands.
- Create their own arrangements.
- Post machine-readable music online, share.
- Blend conventional performance with algorithmic composition, new sounds, new music.
- Robot performers.
- Eventually ... new art forms
- Think of the electric guitar, drum machine in music, camera in visual art, ...

© 2013 Roger B. Dannenberg

**38**

## Another Example



## Conclusion

- Music Understanding and Human Computer Music Performance will *enrich musical experiences* for millions of people, including both amateurs and professionals.
- If we build computers that can perform popular music interactively with intelligence, great music will be made. *That is the future of music performance.*