Some challenges in (creation and) co-creation of music with computers

Geber Ramalho
Universidade Federal de Pernambuco
“If I had an hour to solve a problem, I'd spend 55 minutes thinking about the problem and 5 minutes thinking about solutions.”

Albert Einstein
Foreword

• 2-part talk in a design perspective
  • Why should we explore more co-creation than fully automated creation in music? Musical challenges!
  • If we move towards co-creation, who does what, when, and where?

• Premisse
  • In the automated generation of X (music, text, images, game design...), it would be important to discuss more about X, not just about algorithms!
Why should we explore more co-creation than fully automated creation in music?
Some philosophical reasons... (from François Pachet)

• What is the point to create a technology that does not impact the process of human music creation? Sense of belonging ...
  • Machine learning research that matters for music creation: A case study (Sturm et al)

• A good song is an social object
  • The Hit Song Science – Pachet & Roy

• The hit song paradox
  • “The hit by definition is rare. A machine that generates as many hits you want is a contradiction, by definition” – Pachet’s interview to Folha de São Paulo
And some musical reasons

• Creating good music is f*** hard, even for good musicians

• AI is still a bad composer/improviser
  • Good for creating a “texture” that resembles to a style, but too bad in structuring the discourse/narrative (local vs. long-term dependencies)

• 4 intertwined activities in composing/improvising
  • Motif development
  • Section structuring
  • Harmony progression creation
  • Dramatic curve development
Garota de Ipanema (Jobim/Moraes)

**Remarks**
- **Motif 1:**
  - simple, composed by a cell (sub-motif), with rhythmic elements of samba
  - 3 occurrences with bigger variation in the 3rd, 2nd occurrence without transposition over different chord
- Motif 1 and 2 are very rhythmically (and harmonically) different
- Motif 3 has a beautiful octave jump....

**Questions**
- How Jobim came up with these motifs?
- Why did he developed them such a way?
- How did he decide how many times to repeat/variate them?

Motif development is somehow part of structuring but also provides other qualities (cohesion?)!
Garota de Ipanema (Jobim/Moraes)

A x2
16 bars

B
16 bars

A
8 bars

Remarks
- A classical AABA, but not with 32 bars. It is a 40-bar song! Why?
- B part has 2 parts: a 12-bar one and a 4-bar one: why?
- Transition A to B parts. Contrast both in lyrics, harmony and melody
Garota de Ipanema (Jobim/Moraes)

Remarks
- A part is just a walk around the corner: I–V7/V–IIIm7–V7–I (despite the turnaround! Am7 Ab7 Db7M)
- Whereas B part is a trip + unusual modulations ("Repetition legitimizes!") – or "the equilibrium theory"
“America jazz”

Pery Ribeiro

Jobim/Gilberto

Harmony progression

See Adam Neely’s video for more...
Dramatic curves

- Kurt Vonnegut, 
  Shape of Stories
• “Fortune” = Dramatic tension? Music tension?...

• How? Lots of tools
  • Higher intensity (loudness)
  • Rubato and other expressiveness tools
  • Higher melodic complexity and/or range
  • Higher rhythmic complexity and/or density
  • Higher harmonic complexity
  • Orchestration (timbre)
  • ....

D. Herremans and E. Chew, “Tension ribbons: Quantifying and visualising tonal tension”, 2016
10 guitarists improvise on the same 8-bar progression - Paul Davids

- @Tim Pierce Guitar: https://youtu.be/Tl8CaZVMQzw?t=146
- @Rhett Shull: https://youtu.be/Tl8CaZVMQzw?t=601
- @Rick Beato: https://youtu.be/Tl8CaZVMQzw?t=430
- @Peter Honoré: https://youtu.be/Tl8CaZVMQzw?t=1102
- @Daniel TheGigRig: https://youtu.be/Tl8CaZVMQzw?t=1288
- @Music is Win: https://youtu.be/Tl8CaZVMQzw?t=933
- @Ben Levin: https://youtu.be/Tl8CaZVMQzw?t=1453
- @samuraiguitarist: https://youtu.be/Tl8CaZVMQzw?t=292
- @Chris Buck: https://youtu.be/Tl8CaZVMQzw?t=1622
- @Pete Thorn: https://youtu.be/Tl8CaZVMQzw?t=770
Generating music is somehow like writing a book!

• We cannot do this only by choosing good/right words!
  • Genre, Context (time, geographic, cultural), Theme (and possibly a message), Characters, Plot

• So, the problem of generating good music is probably badly put for lots of algorithms in the literature
  • It is not about only local decisions in a single dimension!
  • But about how to articulate different dimensions in a narrative that conveys an engaging listener experience
Either by some philosophical sanity or by musical humility ...

the problem of computer-generated music should be (re)formulated as "how to build an algorithm that helps human composers/improvisers to create better music?"
“I thought that the outputs generated by the folk-rnn v2 model when fed my own musical ideas would be interesting. I was wrong. But while I ended up using the model differently than I assumed when setting out to write my piece, both the capabilities and limitations of the model shaped the piece and the composition process that led to it. ‘Bastard Tunes’, a four-movement ensemble piece (fl./pic., cl./bass cl., perc., piano, vl., db.), could not have been produced by either myself or folk-rnn without the other”

- Bob Sturm
Co-creation ok, but who does what, when, where? And who decides?
Design space

• The space representing the possible design choices

• Methods using the design space
  • Morphological analysis (Fritz Zwicky 67)
  • Morphological box

• Steps
  • Identify the important dimensions of the problem
  • Determine dimensions range (domains of the variables)
  • *(Apply Cross Consistency Assessment Matrix to eliminate inconsistent options)*
  • Explore possible options
Different applications in product design and HCI

A Morphological Analysis of the Design Space of Input Devices
STUART K. CARD, JOCK D. MACKINLAY, and GEORGE G. ROBERTSON
Xerox Palo Alto Research Center

The Design Space of 3D Printable Interactivity
RAFAEL BALLAGAS and SARTHAK GHOSH, HP Inc., USA
JAMES LANDAY, Stanford University, USA

Design Space for Graph Neural Networks
Jiaxuan You, Rex Ying, Jure Leskovec
Department of Computer Science, Stanford University
{jiaxuan, rexy, jure}@cs.stanford.edu

A Design Space for Gaze Interaction on Head-Mounted Displays
Teresa Hirtle
Institute of Media Informatics, Ulm University, Germany
Jan Gugenheimer
Institute of Media Informatics, Ulm University, Germany
Florian Geiselhart
Institute of Media Informatics, Ulm University, Germany
Andreas Bulling
Institute for Visualisation and Interactive Systems, University of Stuttgart, Germany
Enrico Rukzio
Institute of Media Informatics, Ulm University, Germany

The Design Space of Wireless Sensor Networks
KAY RÖMER and FRIEDEMANN MATTEN, ETH ZURICH

A Design Space of Visualization Tasks
Hans-Jörg Schulz, Thomas Nocke, Magnus Heitzler, and Heidrun Schumann
Useful for

• Systematically evaluating and discovering possible design alternatives for a solution, enhancing creativity
• Choosing a possible instance
• Figuring out holes (in the state of the art)
The Design Space of Music Co-creation (from the interaction perspective)

with Anderson Rufino
What is a co-creation interactive system?

**Creativity Support Tool**
Track history, simulate and explore alternatives to support a creative person

**Generative Systems**
Programs that automatically generate novel, surprising, and valuable creative products

**Computer Colleagues**
Co-creative agents collaborate with humans in *continuous* real-time improvisation to enrich the creative process

Continuous interactions between the machine and the musician that influences each other next perceptions, choices and actions

*Our temporary definition*

Davis, N. M. et al. An Enactive Model of Creativity for Computational Collaboration and Co-creation. 2015
Dimensions and range on music co-creation

1. Musician’s contribution during interaction: what is provided by the human musician?
   • Musical elements
     • Motif
     • Harmonization and counterpoint
     • Harmony
     • Melody
     • Rhythm (only)
     • Sections structure (explicitly)
     • A dramatic curve
   • Dataset (of any musical material)
   • Guidelines
     • Constraints, rules, preferences...
   • Evaluation feedback

2. Machine’s contribution during interaction: what is provided by the machine?
   • Musical elements
     • Motif
     • Harmonization and counterpoint
     • Harmony
     • Melody
     • Rhythm only
     • Sections structure (explicitly)
     • A dramatic curve
   • Guidelines
     • Constraints, rules, preferences...
   • Evaluation feedback
Dimensions and range on music co-creation

3 Musician’s temporal extension: what is the extension of what will be provided by the musician?
   • An excerpt
   • The entire song/piece

4 Machine’s temporal extension: what is the extension of what will be generated by the machine?
   • An excerpt
   • The entire song/piece
Dimensions and range on music co-creation

5. The machine’s role during the interaction with the musician
   - Creation (how)*
     - No explicit musician’s guidelines (ex nihilo or from dataset only)
     - Under musician’s explicit guidelines
     - Given an musician’s example
   - Refinement/improvement

6. The interacting (temporal) protocol (and musician’s autonomy)
   - Single (one-shot) interaction
   - Turning taking
     - In real time
     - Asynchronous/offline
   - Parallel real-time mutual influence

* Should we detail the task? Composition, improvisation, accompaniment
Other possible dimensions not considered

• **Employed generative technique**
  - Markov Chains
  - Formal grammars
  - Rules/constraint based systems
  - Deep learning
  - Evolutionary algorithms
  - Chaos/self-similarity
  - Agent-based systems

• **Representation (musician and machine)**
  - Signal
  - Symbolic

• **Triggers for machine and/or musician action**
  - Fixed duration
  - Silence
  - Button...

• **Types of input**
  - Conventional vs. New midia

• **Activity nature**
  - Composition: criação s/ tempo real
  - Improvisation: criação em tempo real
  - Acompaniment: criação c/ ou s/ TR
  - **Interpretation:** ??????

• **Musician autonomy**
  - Human in the Loop (musician accepts or not each machine proposition)
  - “Human on the Loop” (musician may interfere but has no final decision)

*Subsumed by the interaction protocol dimension*
Some current systems (from Carnovalini and Rodà 2020)

- Music generations systems found: 41
  - Directly cited. No snowballing

Interactive co-creation

- 28; 68%
- 13; 32%
For the non-interactive systems

• Only one example of refinement (NEvMuse), the others on creation

• And concerning machine’s contribution, too many in melody
The interactive ones

<table>
<thead>
<tr>
<th>System/Authors</th>
<th>Designed for interaction</th>
<th>Musician's contribution</th>
<th>Musician's temporal extension</th>
<th>Machine role</th>
<th>Machine's contribution</th>
<th>Machine's Temporal extension</th>
<th>Interaction protocol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuator</td>
<td>yes</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Create</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Real time Turning taking</td>
</tr>
<tr>
<td>Pachet, 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kinetic Engine</td>
<td>Yes</td>
<td>Dataset</td>
<td>Excerpt</td>
<td>Create</td>
<td>Harmony, melody, rhythm</td>
<td>Excerpt</td>
<td>Asynchronous turning taking / Parallel real-time</td>
</tr>
<tr>
<td>Eigenfeidt and Pasquier, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monterey Mirror</td>
<td>yes</td>
<td>Melody or rhythmic pattern (via MIDI)</td>
<td>Excerpt</td>
<td>Create</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Real-time turning taking</td>
</tr>
<tr>
<td>Manaris, Hughes and Vassilandonakis (2011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow Machine</td>
<td>yes</td>
<td>Dataset, melody, harmony, guidelines</td>
<td>Excerpt</td>
<td>Create</td>
<td>Melody, Harmony, Rhythm, Orchestration</td>
<td>Excerpt / Entire song</td>
<td>Asynchronous Turning taking</td>
</tr>
<tr>
<td>Pachet, Roy and Carré (2021)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folk-RNN</td>
<td>Yes</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Create</td>
<td>Melody/Motif variations</td>
<td>Excerpt</td>
<td>Asynchronous Turning taking</td>
</tr>
<tr>
<td>Bob Sturm et al. (2019)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWARMUSIC</td>
<td>yes</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Create</td>
<td>Melody</td>
<td>Excerpt</td>
<td>Real-time turning taking</td>
</tr>
<tr>
<td>Blackwell and Bentley 2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension</td>
<td>Proportion</td>
<td>Not appeared</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **Musician's contribution**| Melody 50% / Harmony 6%  
Dataset 11%  
Eval feedback 11%  
Rhythm 11%  
Guidelines 11% | Motif  
Harmoniz. and counterpoint  
Sections structure  
A dramatic curve |
| **Musician's temporal extension** | Excerpt 79%  
Entire 21% | - |
| **Machine role**           | 100% Creation                                                             | - |
| **Machine's contribution**  | Melody 53%  
Harmony 26%  
Rhythm 11%  
Orchestration 5%  
Motif 5% | Motif  
Guidelines  
Sections structure  
A dramatic curve |
| **Machine's Temporal extension** | Excerpt 71%  
Entire 29% | - |
| **Interaction protocol**   | Real-time turning taking 38%  
Asyncr. turning taking 31%  
Parallel real-time 31% | - |
In short... there are many unexplored or little explored possibilities in designing interactive co-creation systems.