Music Education Framework based on Generative Neural Networks

musicai.citi.sinica.edu.tw y www.dreamstime.com
About me:

Argentina

carana@fi.uba.ar

carlosarana@uca.edu.ar
Carlos "Charly" Arana is a guitarist, producer and researcher specialized in Latin American rhythms and in the application of machine learning and AI techniques to music. As a guitarist, arranger and musical director he has worked and recorded with a multitude of artists from the international scene, both in Argentina, Brazil and the United States, being worthy of mention that he was a band member of legendary Bossa Nova singer Maria Creuza. From 2004 to date, he has edited four books for the two most important music publishers, Hal Leonard and Warner Bros. Publications (Alfred Publications). For his studies and research on machine learning and artificial intelligence he has been invited as a speaker in congresses and seminars at some of the most prestigious technology universities, such as the Massachusetts Institute of Technology (MIT) and UC Berkeley.
General Inspiration:

Formal standardized music education does not encourage "audiation".

Concept of Audiation
Audiation is the foundation of musicianship. It takes place when we hear and comprehend music for which the sound is no longer or may never have been present. One may audiate when listening to music, performing from notation, playing “by ear,” improvising or composing.

Audiation is not the same as aural perception, which occurs simultaneously with the reception of sound through the ears. It is a cognitive process by which the brain gives meaning to musical sounds.

Audiation is the musical equivalent of thinking in language. !!!

The term “audiation” was coined around 1976 by music professor, researcher, theorist, and author Dr. Edwin Gordon (1927-2015)
The most important musicians in popular music learned naturally from **audiation**:
https://www.sightreadingfactory.com/practice/sr/play?keySigs=C%20Major&timeSigs=4/4&levelId=s~3&mediumId=guitar
Al menos los estudios de Carulli y de Leavitt eran súper musicales....
Use of RNN for Musical Training
Ciertas notas/silencios que no reconoce → control de consistencia

Cambio de Tonalidad
What is music21?

Music21 is a Python-based toolkit for computer-aided musicology.

People use music21 to answer questions from musicology using computers, to study large datasets of music, to generate musical examples, to teach fundamentals of music theory, to edit musical notation, study music and the
Bars creation

```

def crea_compases(song):
    """Crea compases a partir del Objeto 'metric.TimeSignature'
    por music21"
    param song :(m21 stream)
    ...  
    # Crea compases
    for event in song.flat:
        if isinstance(event, meter.TimeSignature):
            song.makeMeasures(inplace=True)

    return song
```

```
In [424]: s.flat.show(‘text’)
```

```
In [428]: s.show(‘text’)
```
Key invariance

```python
# arma la primera parte
for i in range(split_1, len(song.getElementsByTagName('Measure'))):
    song_split_2.insert(song.getElementsByTagName('Measure')[i])

# agrega las dos canciones spliteadas
songs.append(song_split_1)
nombre_1 = nombreTema + '_parte 1'
nombres.append(nombre_1)
print('Agregó tema: ' + nombre_1)

songs.append(song_split_2)
nombre_2 = nombreTema + '_parte 2'
nombres.append(nombre_2)
print('Agregó tema: ' + nombre_2)

else:
songs.append(song)
nombres.append(os.path.splitext(file)[0])

return songs, nombres
```
176 canciones (sólo en 4/4)

8183 compases
1) **Objective:** Generation of Tango melodies, to be executed in a sheet music editing software, a music education app or an audio player.

2) **Representation:** symbolic, Note Hold-End representation in *DeepBach* style, with note name in MIDI format, "r" symbol for rest, and a "_" sign, indicating the addition of a sixteenth note (0.25 quarter note, "quarterLength") to the previous note. We would end up with three symbols for staff events: note, rest and extension of a sixteenth note.

3) **Architecture:** LSTM and GAN.

4) **Challenge:** To simulate real-life musical situations and to exercise listening skills.

5) **Strategy:** first start with recurrent networks of the LSTM type. To these will be added methods of sampling, such as temperature and k and p sampling. I will then experiment with GANs conditioned on the underlying chord progression.
Melody generated by the model
**Future Work**

**Part A:** improvement of monophonic melodies
1. Curation of the corpus, incorporating chords
2. Testing alternative recurrent network architectures (LSTM) with Attention (Transformers)

**Part B:**
1. Conditioning the generation of melodies to a given harmony (underlying chord progression).
   Analysis of conditional adversarial generative architecture.

---

**Conditional Generative Adversarial Nets**

Mehdi Mirza  
Département d’informatique et de recherche opérationnelle  
Université de Montréal  
Montréal, QC H3C 3J7  
mirzamom@iro.umontreal.ca

Simon Osindero  
Flickr / Yahoo Inc.  
San Francisco, CA 94103  
osindero@yahoo-inc.com
MIDINET: A CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORK FOR SYMBOLIC-DOMAIN MUSIC GENERATION

Li-Chia Yang, Szu-Yu Chou, Yi-Hsuan Yang
Research Center for IT innovation, Academia Sinica, Taipei, Taiwan
{richard40148, fearofchou, yang}@citi.sinica.edu.tw

https://arxiv.org/abs/1703.10847

Figure 1. System diagram of the proposed MidiNet model for symbolic-domain music generation.

<table>
<thead>
<tr>
<th>Table 2. 13-dimensional chord representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>dimensions 1–12</td>
</tr>
<tr>
<td>major</td>
</tr>
<tr>
<td>C, C#, D, D#, E, F, F#, G, G#, A, A#, B</td>
</tr>
<tr>
<td>minor</td>
</tr>
<tr>
<td>A, A#, B, C, C#, D, D#, E, F, F#, G, G#</td>
</tr>
</tbody>
</table>

Conditional GAN (CGAN)

Generator now generate samples based on some conditions

Discriminator now examine whether a pair \((x, y)\) or \((G(z), y)\) is real or not

Evaluation of generative models in music

Objective evaluation in music generation:

(i) probabilistic measures with-out musical domain knowledge,  
(ii) task/model specific metrics  
(iii) metrics using general musical domain knowledge.
My research

Song empirical probability mass distributions of new metrics:

- Alterations per bar
- Note durations per bar
- Changes in melodic direction
- Melodic range

\[ P_X(x) \]

\[ \frac{1}{2} \]

\[ \frac{1}{4} \]

\[ x \]
Statistical similarity between two empirical mass probability distributions

Kullback-Leibler

\[ D_{KL}(P||Q) = \sum_{x \in X} P(x) \ln \frac{P(x)}{Q(x)} = E_p \ln \frac{P(X)}{Q(X)} \]

Shannon-Jensen

\[ JSD(P||Q) = \frac{1}{2} D(P||M) + \frac{1}{2} D(Q||M) \]
\[ JSD(P||Q) = H \left( \frac{P + Q}{2} \right) - \frac{H(P) + H(Q)}{2} \]