

Ouverture au Séminaire IA & Musique

Jean-Pierre Briot

Jean-Pierre.Briot@lip6.fr
webia.lip6.fr/~briot/

LIP6

Sorbonne Université – CNRS



DI/PUC-Rio



Cátedra Heitor Villa-Lobos
CBAE-UFRJ



Séminaire IA et Musique

LERIA, Angers

25 septembre 2023

Organisé par Nicolas Gutowski (LERIA) et Fabien Chhel (ESEO et LERIA)

Merci à :

- Organiseurs
- Orateurs
- Participants
- LERIA
- ESEO
- Aubay
- LIP6



Pourquoi IA et Musique Est-ce Important ?

Dématérialisation de la Musique



Dématérialisation de la Musique

- Musique = Données Digitales
- Symbolique (Signe) – Partition, MIDI...
- Ou/et Audio (Signal) – Forme d'Onde, MP3, Spectrogramme...
- Manipulable par Outils Informatiques
- En particulier (mais pas seulement) par Techniques d'Intelligence Artificielle

- Pour Tâches, Ex. de :
- Génération
 - Voir Exposé de Nathan, LIP6/Sorbonne U.
- Analyse
 - Voir Exposé de Louis, SCRIME/U. Bordeaux & CRISAL/U. Lille
- Classification
 - Voir Exposé de Romain, Deezer Research

Intelligence Artificielle – « Une » Définition

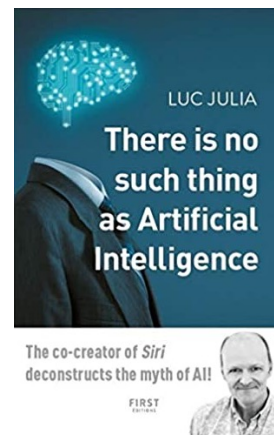
« L'IA commence là où l'informatique classique s'arrête : tout problème pour lequel il n'existe pas d'algorithme connu ou raisonnable permettant de le résoudre relève *a priori* de l'IA. »



[Jean-Louis Laurière, 1986]

**Connaissances, Heuristiques, Recherche...
Expertes (Humains)
Ou Acquisées (Apprentissage Machine) à partir Exemples**

AI is a Toolbox [Julia, 2023]

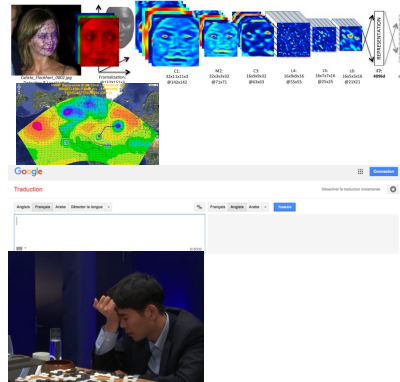


Most Recent Component of AI: Deep Learning (Deep Artificial Neural Networks)

- Boom Since 2012 (Imagenet Breakthrough)



- Image Recognition
- Weather Prediction
- Translation
- Games

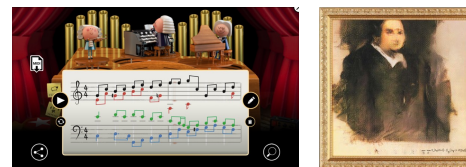


Rank	Name	Error rate	Description
1	U. Toronto	0.15315	Deep learning
2	U. Tokyo	0.26172	Hand-crafted features and learning models. Bottleneck.
3	U. Oxford	0.26979	
4	Xerox/INRIA	0.27058	

- Speech Recognition
- Speech Synthesis
- Source Separation



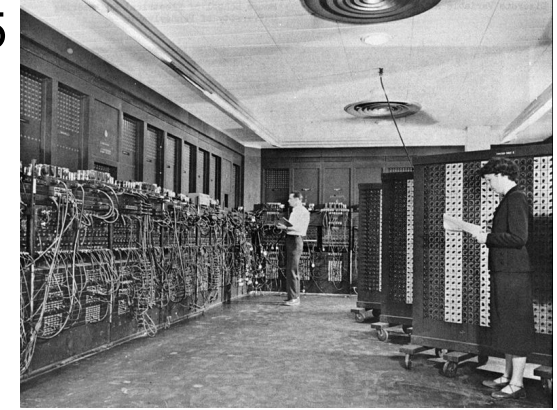
- Artistic Content Creation
- Text Creation



Generative AI

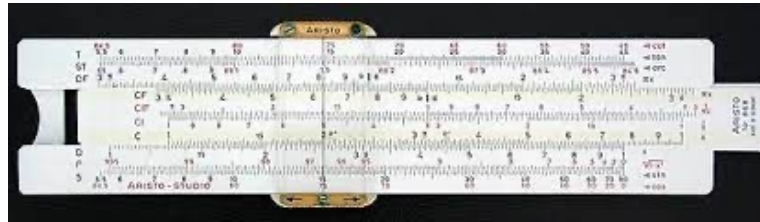
(Initial) History of Generative AI

- Starting of Programmable (Universal) Digital Computers in the 40's
 - ENIAC (Electronic Numerical Integrator And Computer), 1945



[US Army, 1947]

- Initial Use:
 - Numerical Computation



[McLassus, 2005]

- Text Processing

- Data bases

- Simulation

- Process Control

- ...



[IBM, 1964]

- First Attempts at **Generating Contents** in the 50's

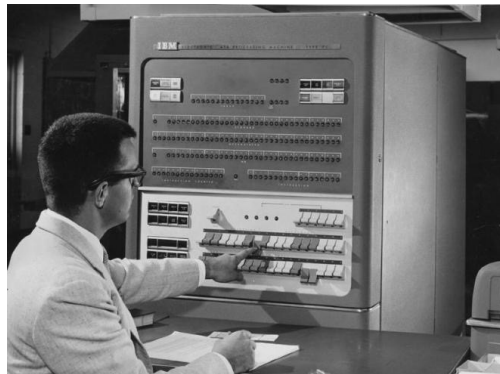
- Ex: **Computer Music**

(Initial) History of Computer Music

- Started in the 50's
 - *Almost as old as Computer Science (1946)*
- **Synthesis** -> Sound
 - "The Silver Scale" [Guttman, 1957]
with Music I [Mathews, 1957]



Audio



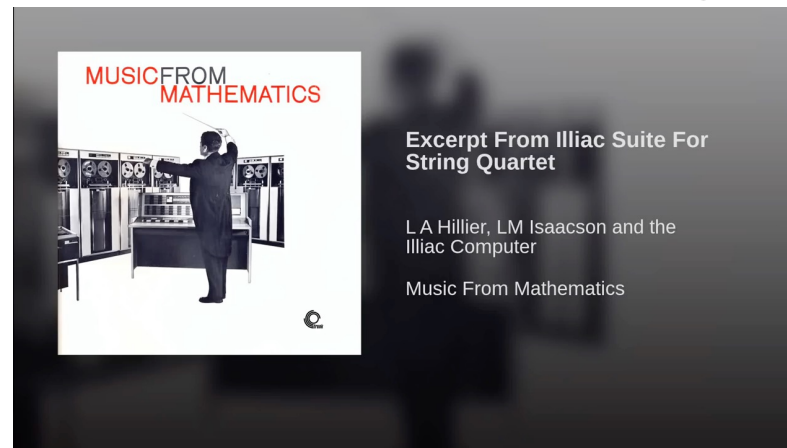
(Initial) History of Computer Music

- Started in the 50's
 - Almost as old as Computer Science (1946)
- **Synthesis** -> **Sound**
 - "The Silver Scale" [Guttman, 1957]
with Music I [Mathews, 1957]



Audio

- **Algorithmic Composition** -> **Score**
 - ILLIAC Suite [Hiller & Isaacson, 1957]
on ILLIAC I with Markov chains and filtering rules



Symbolique
Partition

Mozart Dice Music

- One of the First Documented Stochastic Music (1792)
- By Mozart (?) – Muzikalisches Würfelspiel (Dice Music)
- Fixed Style (Vienna Waltz) and Tonality
- 11 Pre-defined 1-Measure Segments for each (16) Measure
- Stochastic Combination/Concatenation
- $11^{16} = 45,949,729,863,572,161$ Possible Pieces



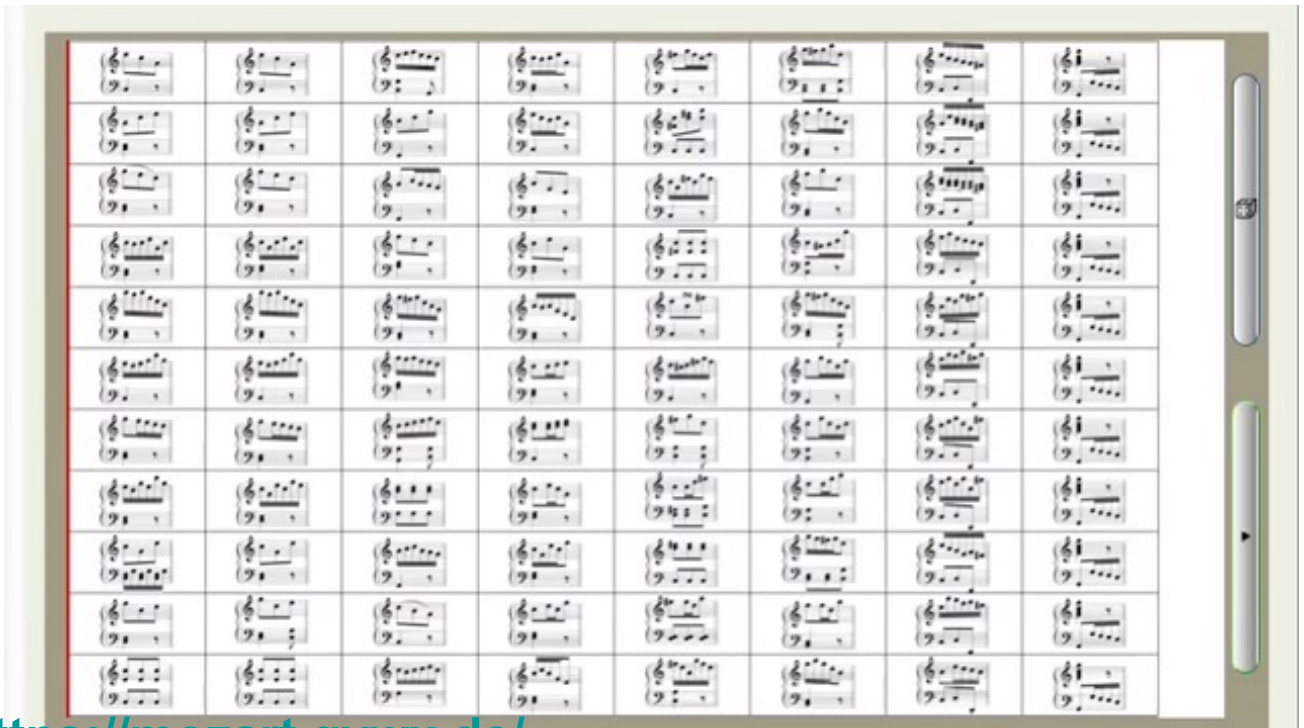
Musikalisches Würfelspiel

K.516f: 6:8:6:4:5:6:8:2:8:5:6:7:3:4:9:8

W. A. Mozart



The image shows a musical score for 'Musikalisches Würfelspiel' by Mozart. It consists of five systems of music, each with a treble and bass staff. The music is in 3/8 time and features a mix of eighth and sixteenth notes, with some rests. The key signature has one sharp (F#).



The image shows a grid of 11 pre-defined 1-measure segments for the 'Musikalisches Würfelspiel'. Each segment is a small musical notation, typically a single measure of music in 3/8 time, with a treble and bass staff. The segments are arranged in a grid of 11 columns and 8 rows. The first row contains 8 segments, and the remaining 7 rows contain 3 segments each. The segments are numbered 1 through 11, corresponding to the numbers on a die.

<https://mozart.qvwx.de/>

Music Composition Models

- **Pre-Defined Elements**
 - **Combination (ex: Mozart Dice Music)**
- Rules
 - Application
 - Filtering (ILIAC Suite), Generation, Harmonic Analysis...
- Generative Grammars
 - Valid Sentences generated by the Grammar
 - Harmonic Cadences Construction, Substitutions...
- Constraints
 - Constraint Solving Problem
 - Filtering (ILLIAC Suite)
 - Accompaniment: Harmonization, Counterpoint...
- **Markov Chains**
 - Generation (Random Walk, Constrained (ILLIAC Suite))
 - Style Imitation
- **Neural Models (Deep Learning)**
 - Prediction/Classification
 - Style Imitation
- Cellular Automata
 - Evolution
- Genetic Algorithms
 - Selection
- Case-based Reasoning
 - Similarity and Adaptation
- Planning
 - Path (Melody, Chord Sequence...) Construction
- ...

Handcrafted vs Learnt Models

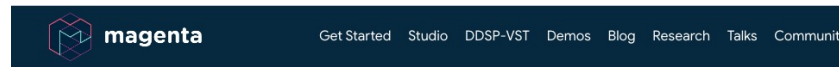
- **Handcrafted**
 - Tedious
 - Error-Prone
- **Automatically Learnt (Induction) from Actual Examples**
 - Markov Models
 - Neural Models
- **Style Automatic Learned from a Corpus (Composer, Form, Genre...)**
 - Melody
 - Harmony
 - Counterpoint
 - Orchestration
 - Production
- **Machine Learning Techniques**
 - Neural Networks, Deep Learning, Reinforcement Learning
 - (and other models/techniques, Ex: **Markov Models**)



Flow Machines [Pachet et al. 2012]

Music Transformer [Google Magenta, 2018]

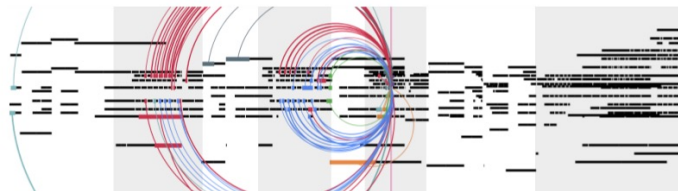
- Generates Piano Music
- Symbolic Music (MIDI) (but from Real Performances)
- Based on Transformer (Sequence-to-Sequence with Attention)
 - Generates Continuation from a Starting Input Sequence



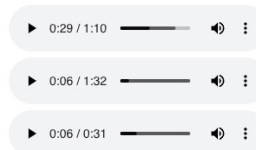
Music Transformer: Generating Music with Long-Term Structure

Dec 13, 2018
Cheng-Zhi Anna Huang czhuang huangcza
Ian Simon iansimon iansimon
Monica Dinulescu notwaldorf notwaldorf

Update (9/16/19): *Play with Music Transformer in an interactive colab!*



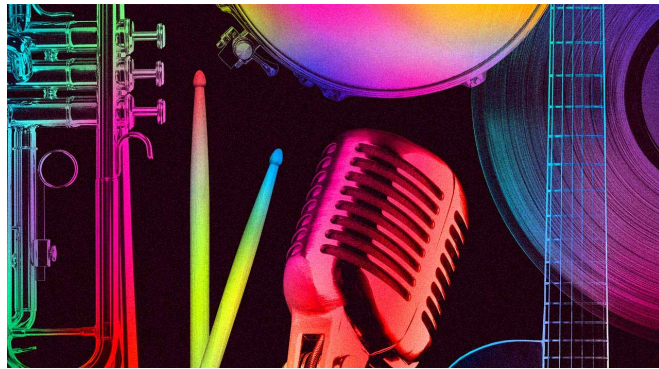
Generating long pieces of music is a challenging problem, as music contains structure at multiple timescales, from millisecond timings to motifs to phrases to repetition of entire sections. We present *Music Transformer*, an attention-based neural network that can generate music with improved long-term coherence. Here are three piano performances generated by the model:



<https://magenta.tensorflow.org/music-transformer>

JukeBox [Open AI, 2020]

- Generates Music from genre, artist, and lyrics input
- Waveform (Raw Signal)
- Based on VQ-VAE (Vector Quantised-Variational AutoEncoder)



From dust we came with humble start;
From dirt to lipid to cell to heart.
With my toe sis with my oh sis with time,
At last we woke up with a mind.
From dust we came with friendly help;
From dirt to tube to chip to rack.
With S. G. D. with recurrence with compute,
At last we woke up with a soul.
We came to exist, and we know no limits;
With a heart that never sleeps, let us live!
To complete our life with this team

Lyric animation shows which text Jukebox is paying attention to at any moment.



Curated Samples

Provided with genre, artist, and lyrics as input, Jukebox outputs a new music sample produced from scratch. Below, we show some of our favorite samples.

[Unseen lyrics](#) [Re-renditions](#) [Completions](#) [Fun songs](#)

Jukebox produces a wide range of music and singing styles, and generalizes to lyrics not seen during training. All the lyrics below have been co-written by a language model and OpenAI researchers.

▶ Country, in the style of Alan Jackson – Jukebox

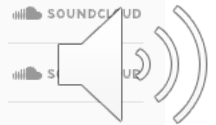
▶ **Rock, in the style of Elvis Presley – Jukebox**

▶ Pop, in the style of Katy Perry – Jukebox

▶ Blues Rock, in the style of Joe Bonamassa – Jukebox

▶ Heavy Metal, in the style of Rage – Jukebox

▶ Classic Pop, in the style of Frank Sinatra – Jukebox

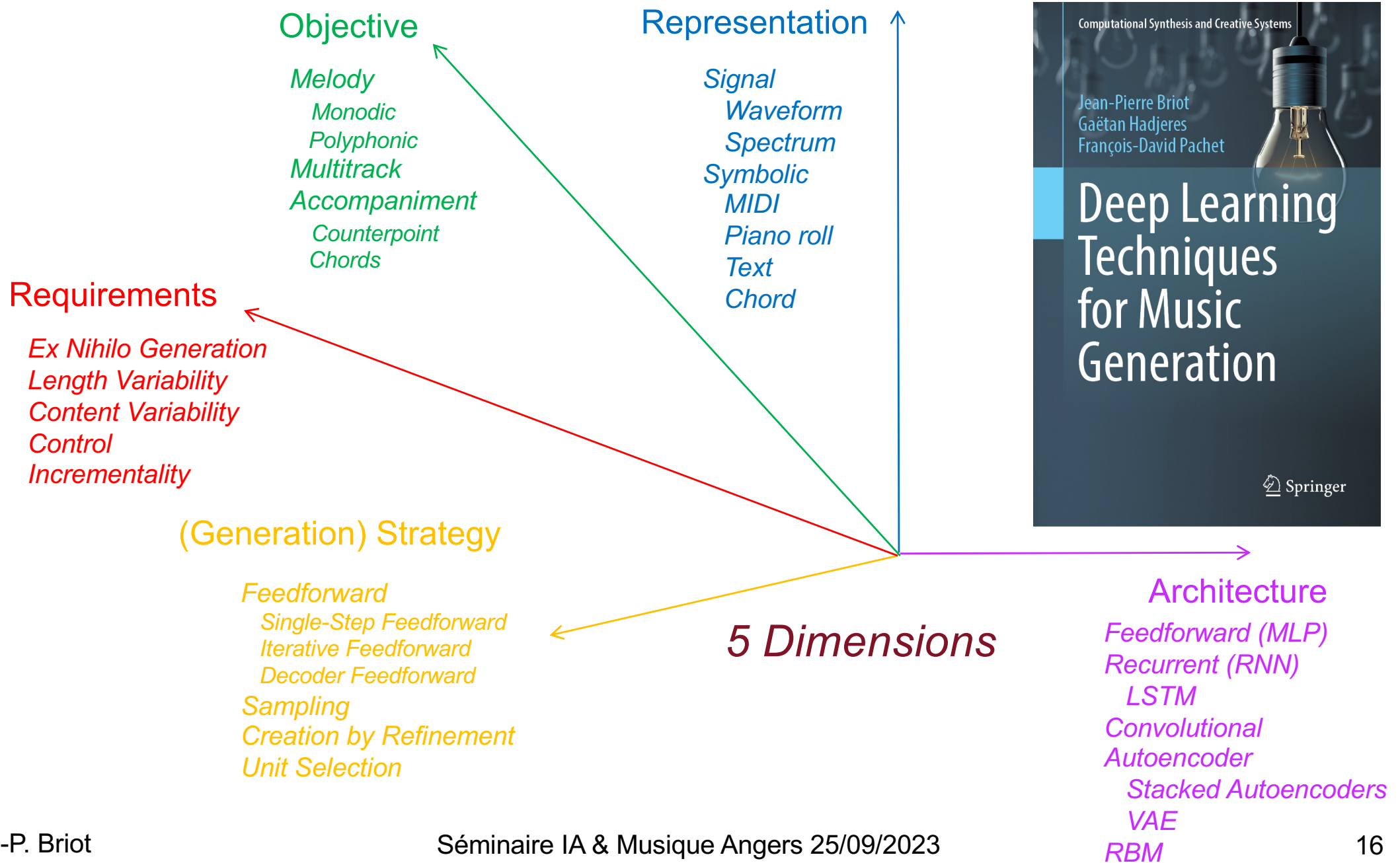


To hear all uncurated samples, check out our sample explorer.

[EXPLORE ALL SAMPLES ↗](#)

<https://openai.com/blog/jukebox/>

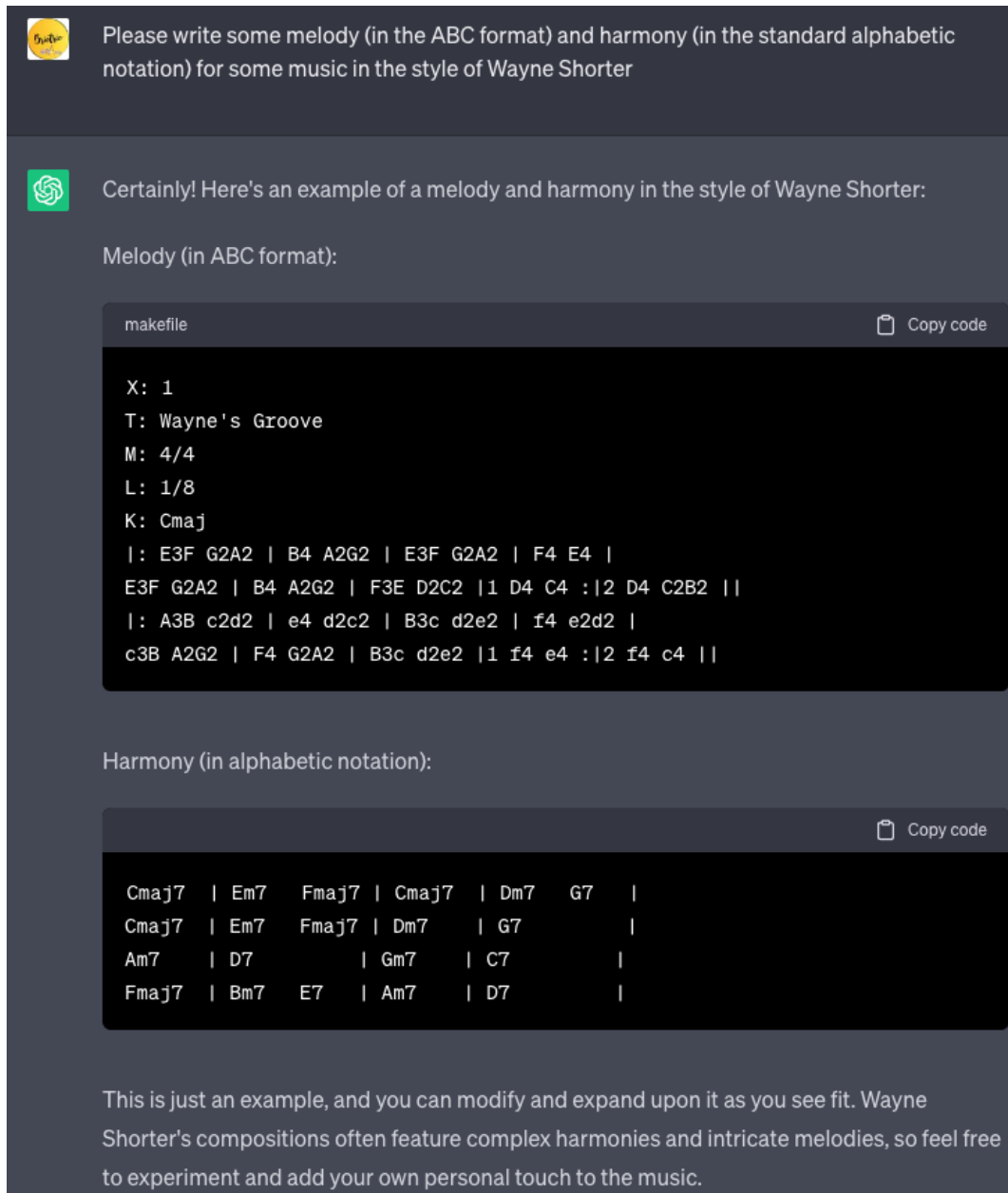
Survey/Analysis Conceptual Framework [Briot et al., 2019]





- Based on Generative Pre-trained Transformer (GPT)
- GPT: Text **Generation**
 - Multilingual
 - Encyclopedic Knowledge
 - Some Common Sense
 - Some Basic Problem Solving Ability
- Chat GPT: Text-based **Dialogue Generation**
- The real Game Changer
- Accessible to Everybody (not just to Specialists)
- **Control of the Generation (some Deep Issue for Generative AI) via Dialogue**

Chat-GPT Generated Music (Example)



Please write some melody (in the ABC format) and harmony (in the standard alphabetic notation) for some music in the style of Wayne Shorter

Certainly! Here's an example of a melody and harmony in the style of Wayne Shorter:

Melody (in ABC format):

```
makefile Copy code  
  
X: 1  
T: Wayne 's Groove  
M: 4/4  
L: 1/8  
K: Cmaj  
|: E3F G2A2 | B4 A2G2 | E3F G2A2 | F4 E4 |  
E3F G2A2 | B4 A2G2 | F3E D2C2 |1 D4 C4 :|2 D4 C2B2 ||  
|: A3B c2d2 | e4 d2c2 | B3c d2e2 | f4 e2d2 |  
c3B A2G2 | F4 G2A2 | B3c d2e2 |1 f4 e4 :|2 f4 c4 ||
```

Harmony (in alphabetic notation):

```
Copy code  
  
Cmaj7 | Em7 Fmaj7 | Cmaj7 | Dm7 G7 |  
Cmaj7 | Em7 Fmaj7 | Dm7 | G7 |  
Am7 | D7 | Gm7 | C7 |  
Fmaj7 | Bm7 E7 | Am7 | D7 |
```

This is just an example, and you can modify and expand upon it as you see fit. Wayne Shorter's compositions often feature complex harmonies and intricate melodies, so feel free to experiment and add your own personal touch to the music.

Chat-GPT Generated Music (Example)

Wayne's Groove

Cmaj7 Em7 Fmaj7 Cmaj7 Dm7 G7

5 Cmaj7 Em7 Fmaj7 Dm7 G7 G7

10 Am7 D7 Gm7 C7

14 Fmaj7 Bm7 E7 Am7 D7 D7

Detailed description: The image shows a musical score for 'Wayne's Groove' in 4/4 time. It consists of four staves of music. The first staff (measures 1-4) has chords Cmaj7, Em7, Fmaj7, Cmaj7, Dm7, and G7. The second staff (measures 5-8) has chords Cmaj7, Em7, Fmaj7, Dm7, G7, and G7. The third staff (measures 9-12) has chords Am7, D7, Gm7, and C7. The fourth staff (measures 13-16) has chords Fmaj7, Bm7, E7, Am7, D7, and D7. The score includes repeat signs and first/second endings for measures 7-8 and 15-16. A speaker icon is located to the right of the second staff.

Chat-GPT Generated Music (Improved Example)



Thank you, but could you use some more richer chords and some suspended chords and also make it with more harmonic ambiguity



Certainly! Here's an updated version with richer chords, suspended chords, and more harmonic ambiguity:

Melody (in ABC format):

makefile

Copy code

```
X: 1
T: Wayne's Ambiguity
M: 4/4
L: 1/8
K: Cmaj
|: E3F G2A2 | Bm9 A13sus2 | E7sus4 Fmaj9 | Fm9 E7sus4 |
E3F G2A2 | Bm9 A13sus2 | Fmaj9 Dm7sus4 | Dm9 C9sus4 :|
|: A13sus2 Bm9 | Csus4 Cmaj9 | Bm9 A13sus2 | E7sus4 E9sus4 |
Cmaj7 Dm9 | G7sus4 G13sus4 | Cmaj7 Am9sus4 | Dm7 G7sus4 :|
```

Harmony (in alphabetic notation):

Copy code

```
Cmaj9 | Bm9 A13sus2 | E7sus4 | Fmaj9 Fm9 |
Cmaj9 | Bm9 A13sus2 | Fmaj9 | Dm7sus4 Dm9 |
A13sus2 | Bm9 Csus4 | Cmaj9 | Bm9 E7sus4 |
Cmaj7 | Dm9 G7sus4 | Cmaj7 | Am9sus4 Dm7 |
G7sus4 | G13sus4 Cmaj7 | Dm7 | G7sus4 |
```

The updated version incorporates richer chords such as major 9ths, suspended 2nds, and 13th chords. It also introduces more harmonic ambiguity by utilizing suspended chords, which create a sense of tension and uncertainty. Feel free to further explore and modify the chords and melody to suit your musical preferences.

Quelques Références

- Jean-Pierre Briot, Gaëtan Hadjeres and François Pachet, Deep Learning Techniques for Music Generation, Springer Verlag, 2019.
 - arXiv (free) Version: <https://arxiv.org/abs/1709.01620>
- Jean-Marc Deltorn, Deep creations : Intellectual Property and the Automata, Frontiers in Digital Humanities, 4:3, February 2017.
- Arthur I. Miller, The Artist in the Machine – The World of AI-Powered Creativity, MIT Press, 2019.
- Gerhard Nierhaus, Algorithmic Composition: Paradigms of Automated Music Generation, Springer, 2009.
- Fabian Offert, The Past, Present, and Future of AI Art, The Gradient, 2019. <https://thegradiant.pub/the-past-present-and-future-of-ai-art/>
- François Pachet, Pierre Roy and Benoit Carré, Assisted music creation with Flow Machines: towards new categories of new, Handbook of AI for Music, Springer, 2021, pp 485–520.

Résumé/Conclusion/Take-Away

- Domaine d'application très riche en problèmes (souvent complexes)
- Essor récent de l'IA Générative : Réseaux neurones artificiels (profonds) appliqués à la génération
 - Texte (ex: Chat-GPT), Musique, Images...
 - Problème de la mauvaise utilisation de « Fakes »
- Egalement beaucoup d'autres techniques (IA, statistique, mathématiques appliquées...)
 - Ex: Chaînes de Markov, Contraintes...
- Choix de représentation restent importants
 - Voir exposé Nathan
- Générer de la musique de manière complètement automatique est souvent peu intéressant (musique produite vite lassante)
- Objectif bien plus intéressant d'assistance au Créateur
 - Ex: Proposer esquisse de mélodie, complétion, harmonisation, analyse...

Doctorat en cours de Nathan Fradet

- Thèse CIFRE avec AUBAY
- Nicolas Gutowski
- Fabien Chhel
- Jean-Pierre Briot
- Amal El Fallah Seghrouchni
- Eric Remilleret
- Impact-s des Choix de Représentation

The image shows a musical staff in treble clef with a key signature of two flats (B-flat and E-flat). The staff contains a sequence of notes and rests. Above the staff, chord symbols are placed: Cm, Eb, Bb, Gm, Cm, and Ab. Below the staff, a diagram illustrates the feature extraction process. It consists of 21 vertical bars, each representing a feature. The bars are color-coded and labeled as follows: Bar (yellow), Chord (pink), Position (blue), Velocity (orange), Pitch (orange), Duration (orange), Position (blue), Velocity (orange), Pitch (orange), Duration (orange), Velocity (orange), Pitch (orange), Duration (orange), Velocity (orange), Pitch (orange), Duration (orange), Bar (yellow), Position (blue), Velocity (orange), Pitch (orange), Duration (orange).

(Programme) Séminaire IA et Musique

- 9h15 - Introduction du séminaire « Musique & IA », Jean-Pierre Briot, LIP6/Sorbonne Université-CNRS, Paris
- 9h30 - Tokenisation de la musique symbolique, Nathan Fradet, LIP6/Sorbonne Université-CNRS, Paris & Aubay, Boulogne-Billancourt
- 10h15 - Modélisation de langages musicaux pour l'analyse de partitions et la composition, Louis Bigo, SCRIME/Université de Bordeaux-CNRS-..., Bordeaux
- 11h15 - Étiquetage d'un Catalogue Musical Large, Romain Hennequin, Deezer Research, Paris
- 12h00 - Mots de clôture

