Deep Learning Techniques for Music Generation
0. General Introduction

Jean-Pierre Briot

Jean-Pierre.Briot@lip6.fr

Laboratoire d'Informatique de Paris 6 (LIP6)
Sorbonne Université – CNRS

Programa de Pós-Graduação em Informática (PPGI)
UNIRIO

Course: Deep Learning for Music Generation

- Survey and Analysis of various deep learning-based music generation systems
- Very active domain
  - Ex: Google Magenta Project
  - Spotify Creator Technology Research Lab (CTRL)
  - Also image and text generation
- Objective: Understand and analyze (classify) various existing approaches and challenges/issues
- Challenge for deep learning, designed for prediction and classification, and not for generation, and furthermore not for creative and controlled generation
- Based on Joint work with Gaëtan Hadjeres (Sony CSL) and François Pachet (Spotify CTRL)

Objective

- Show that Deep Learning Techniques for Music Generation are getting Mature
- Show that Control Issues are Important and Difficult
  - and with some Partial Solutions
- Show that Creativity vs Plagiarism Issues are Important and Difficult
  - and with some Partial Solutions

Outline

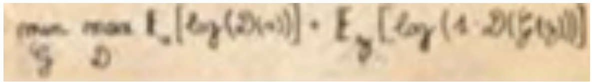
- (Introduction to) Computer Music and Algorithmic Composition
- (Introduction to) Deep Learning and Neural Networks
- Deep Learning for Generating Music
- Representations
  - Audio, Symbolic, Piano roll, MIDI, One-hot…
- Architectures
  - Feedforward, Recurrent, Autoencoder, RBM, GAN, VAE…
- Issues/Strategies
  - Variability/Sampling
  - Control/Input Manipulation (Ex: Deep Dream, Style Transfer, C-RBM)
  - Control/Reinforcement (Ex: RL-Tuner)
  - Creativity (Ex: CAN)
  - Interactivity (Ex: DeepBach)
- Discussion
- Conclusion

This Introduction Outline

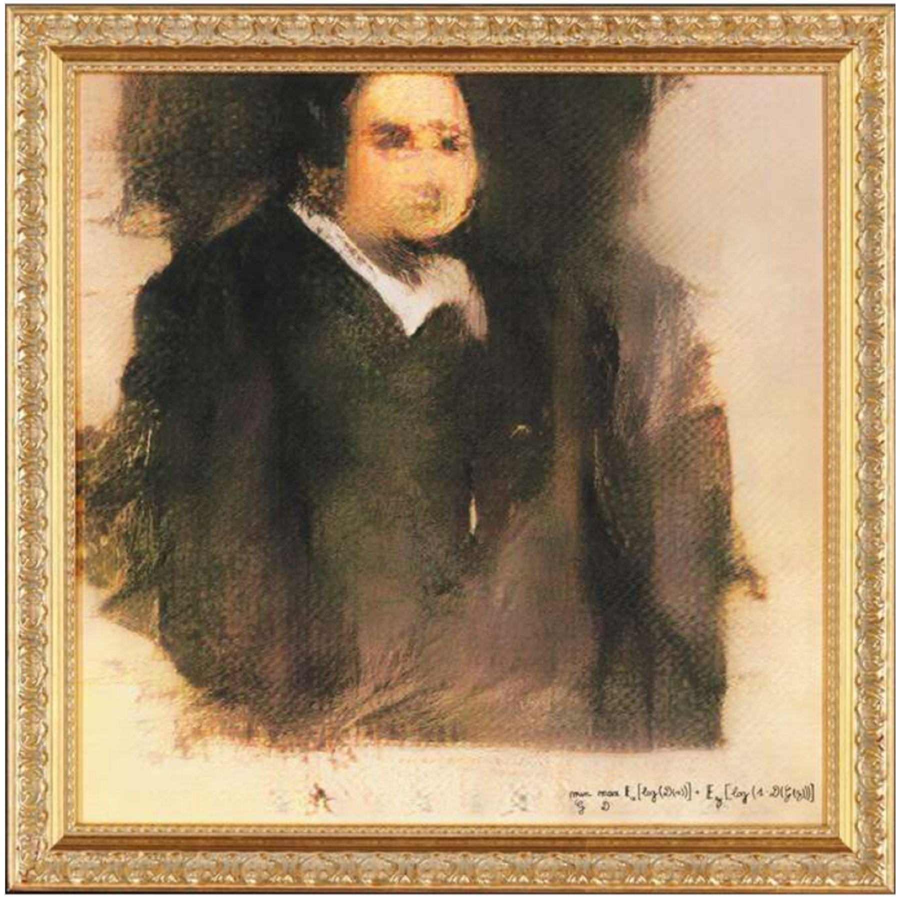
- Recent Artificial Creative Artifacts
- Why Music with Computers?
- Computer Music History
- Autonomous vs Assistance/Interactive Music Making
- AI History
- Autonomous Artificial Intelligence (AI)
  vs Intelligence Augmentation (IA)
- From Perceptron (Neural Networks) to Deep Learning
- Deep Learning Music Creation
- Control
- Conclusion

Recent Creations

Painting

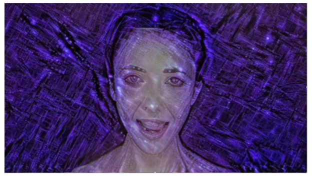
- 26 October 2018, Christie’s Auction, New York, US$ 432 500
- Edmond de Belamy, Obvious (Collective) 

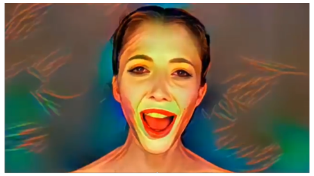
- Created with Deep Learning (GAN)
- Trained with 15 000 paintings (XIV – XX centuries)



Pop Music

- August 2017, Break Free, Taryn Southern
- Music composed by Amper AI
- Video AI Art Created with Deep Dream Generator
- Both Created with Deep Learning

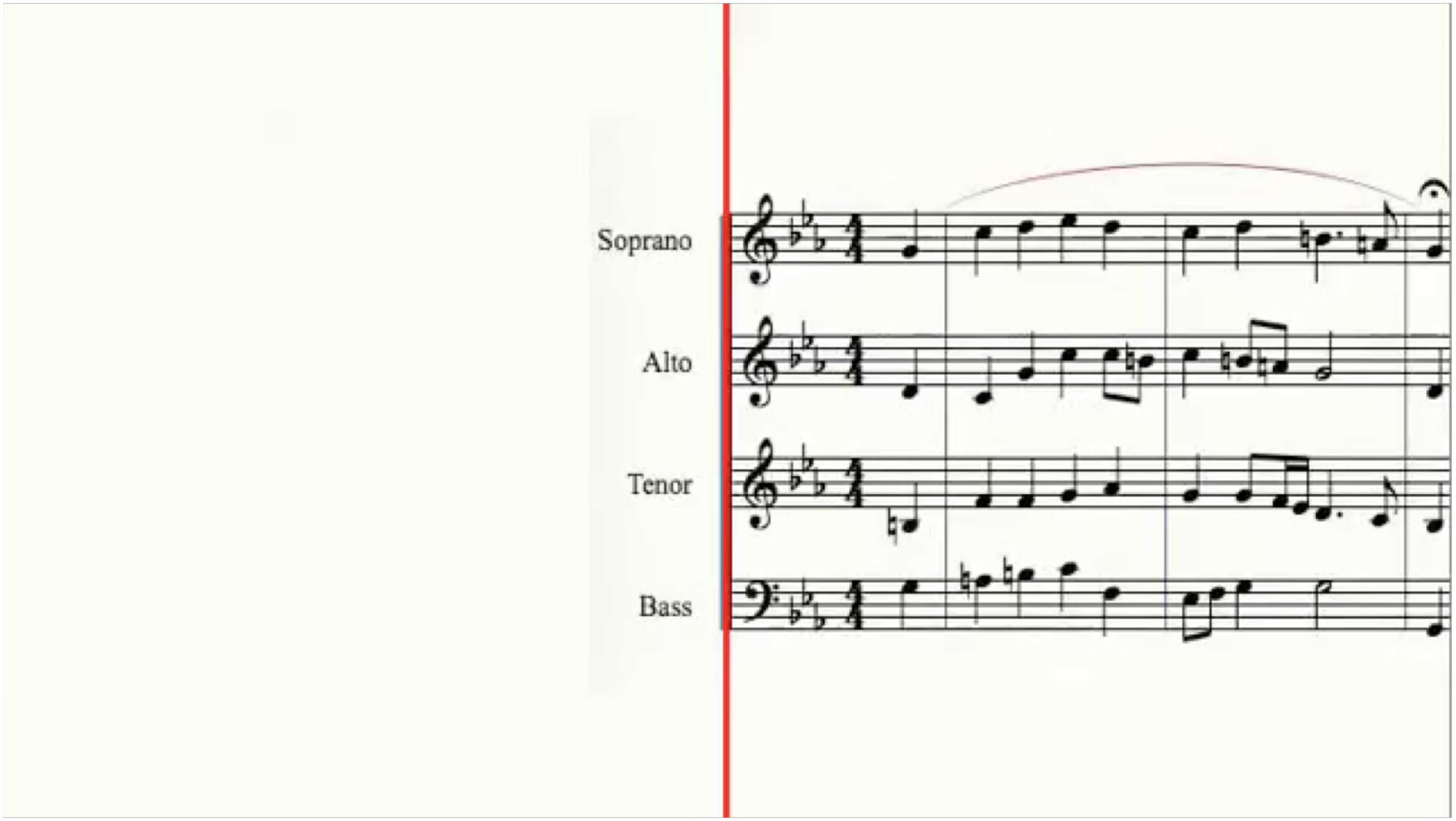




https://www.youtube.com/watch?v=XUs6CznN8pw

Bach Chorales

- December 2016, DeepBach, Gaëtan Hadjeres
- Deep Learning
- Training Set = 352 Chorales



https://www.youtube.com/watch?v=QiBM7-5hA6o

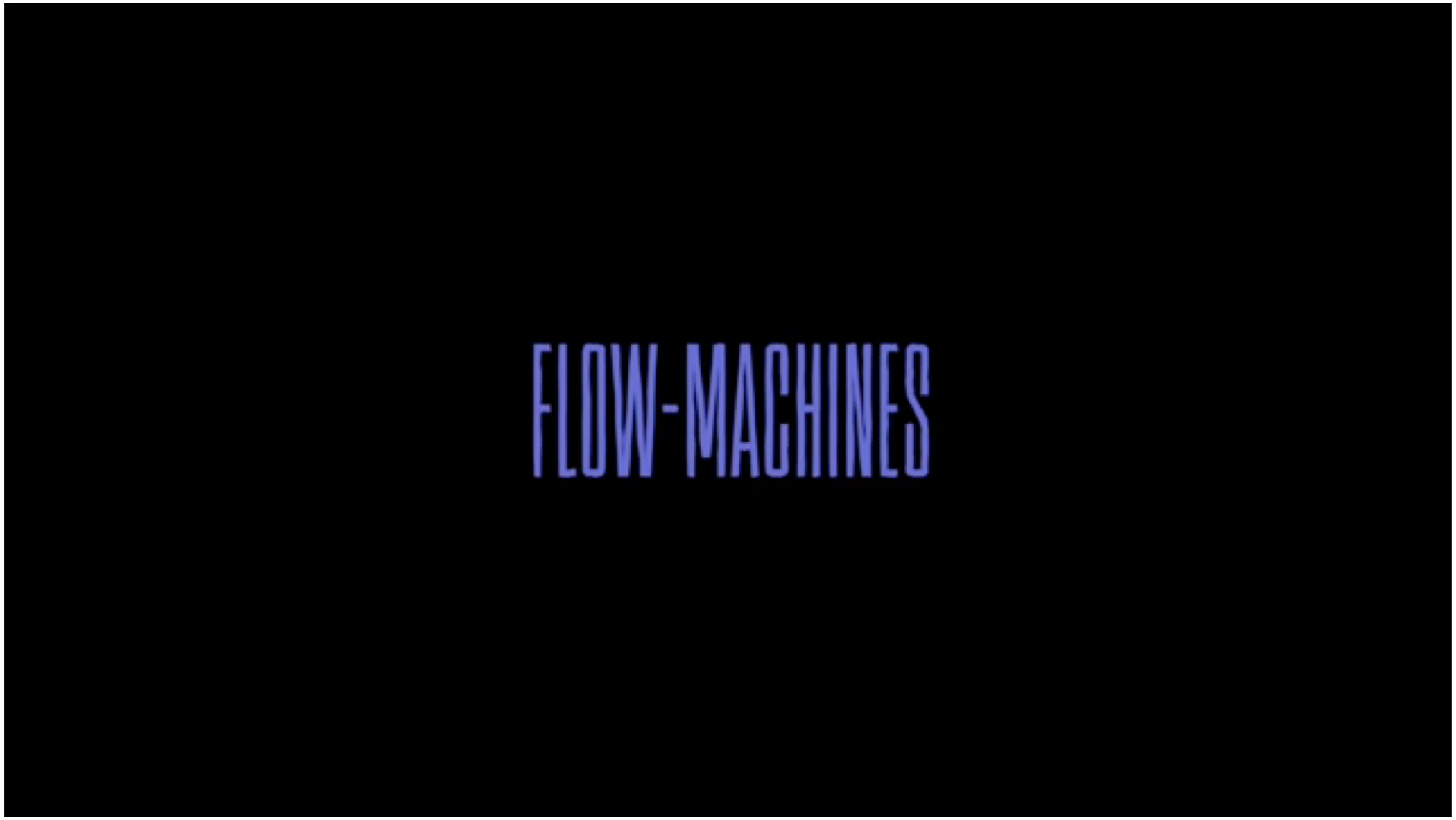
Hello World

- January 2018, Hello World
- Created by Musicians (Musical Direction: Skygge – aka Benoît Carré)
- with FlowComposer [Pachet et al., 2014]
- ERC Project Flow Machines [Pachet et al., 2012-2017]
- Various Techniques (Markov Constraints, Rules, …)

https://www.youtube.com/watch?v=iuWYQe3aGIg

Hello World

- January 2018, Hello World, Flow Records
- Making Off



https://www.youtube.com/watch?v=yxTF-UFvoHU

Motivation

Why Using Computer for Music

- Bad Reasons (Fears)
  - Lead human musicians to unemployment
  - Lower the quality of music ☺

- Good reasons
  - Facilitate storing, indexing, delivering and sharing of music (MIDI, MP3, Spotify…)
  - New instruments and interaction (Synthesizers, Interactive music performances…)
  - New sounds (Synthesizers and Signal processing)
  - Analysis tools and algorithms (Spectrum, Patterns Discovery…)
  - Initiation and Education (Band in the Box, Garage Band…)

- Production
  - Partially automate tasks (Ex: Mixing, etc.)

- Composition, Analysis and Arrangement
  - Algorithmic composition
  - Harmonization
  - Analysis

Why Using Computer for Music

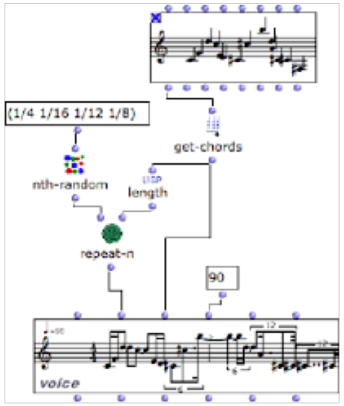
- Vast Associative Memory
  - More systematic than Human memory
- Representation of Musical pieces, Style, Patterns…
- Associations and Correlations
- Knowledge (Theory, Rules, Heuristics…)

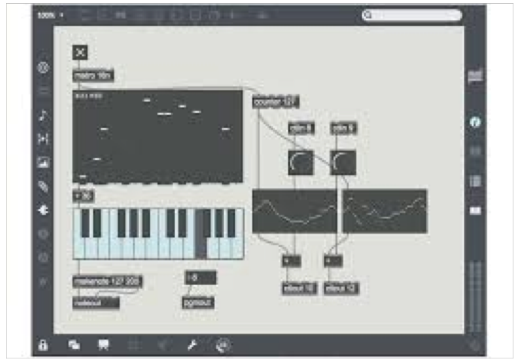
- Can Help Human musicians

- Human musicians rarely compose from scratch – They borrow from others
  - Consciously
    - » Plagiat, Citation…
  - Unconsciously
    - » Influence
  - Recombinations
  - Historical Evolution/Extension
    - » Modal monophonic -> Polyphonic (Counterpoint) -> Tonal Music (Harmony) -> Extended Harmony (Debussy, Jazz…)
  - Ruptures (Dodecaphonism, Free Jazz, Punk…)
    - » Rare and often transient

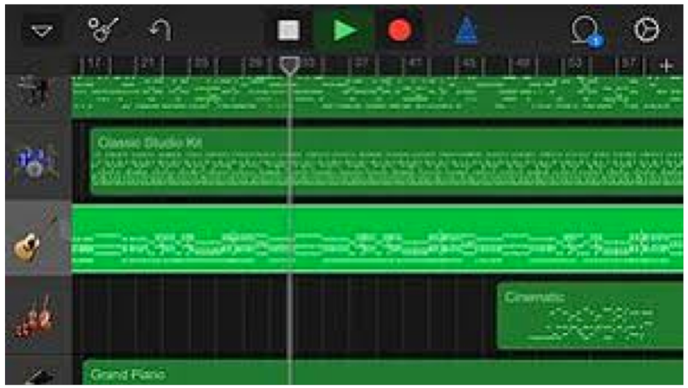
Music Composition/Production Environments

- Expert Musician Algorithmic Composition or/and Real-Time Production Environment
- Ex: OpenMusic, Max/MSP…
  - Musical Concepts
  - Algorithmic Composition
  - Tool-Box
  - Composition vs Interaction
  - Symbolic vs Signal Processing





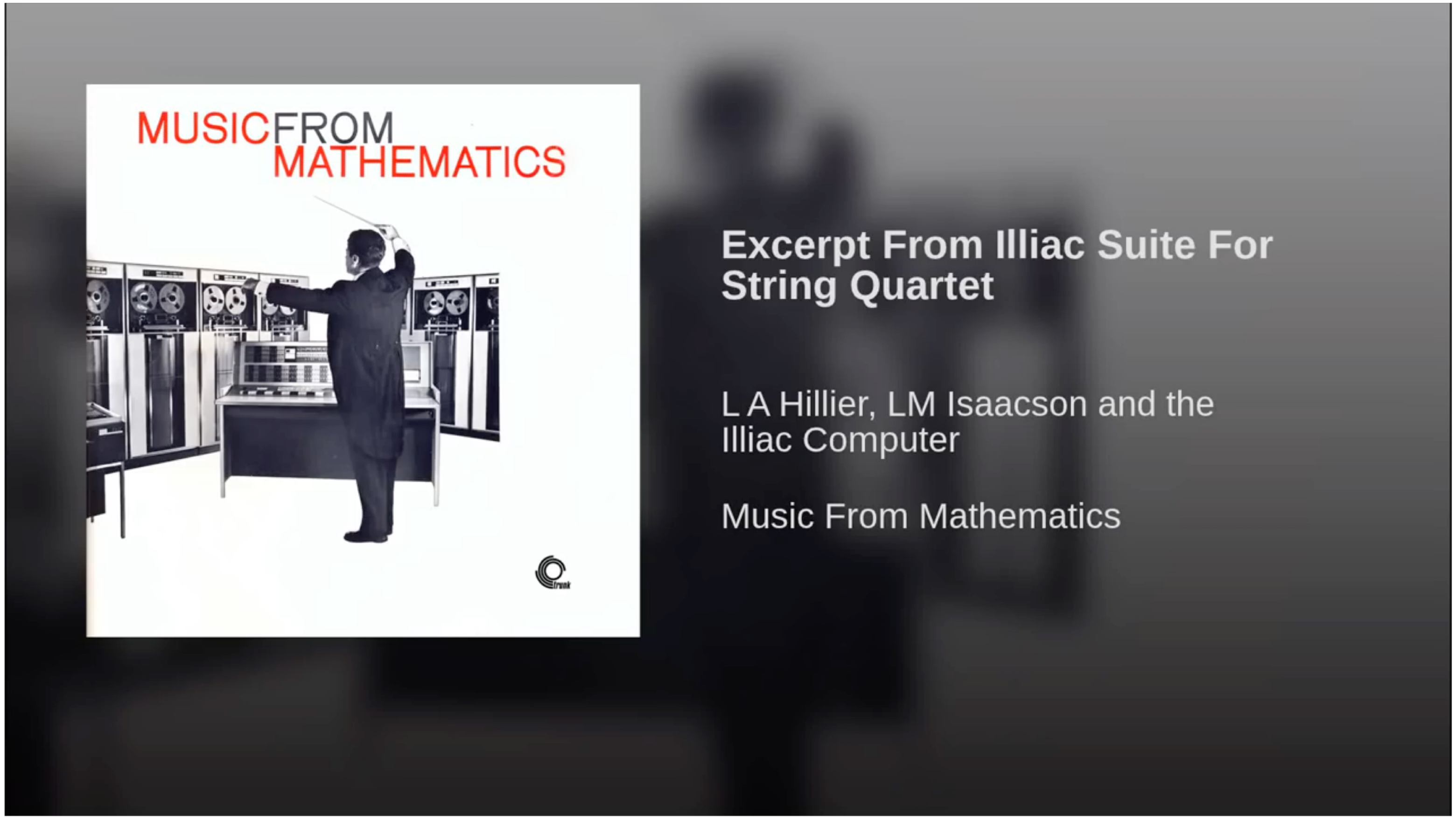
- Average Musician Music Production and Composition Environments
- Ex: GarageBand, Band-in-a-Box…
  - Good Tradeoff
  - Little Musical/Technical Knowledge



History

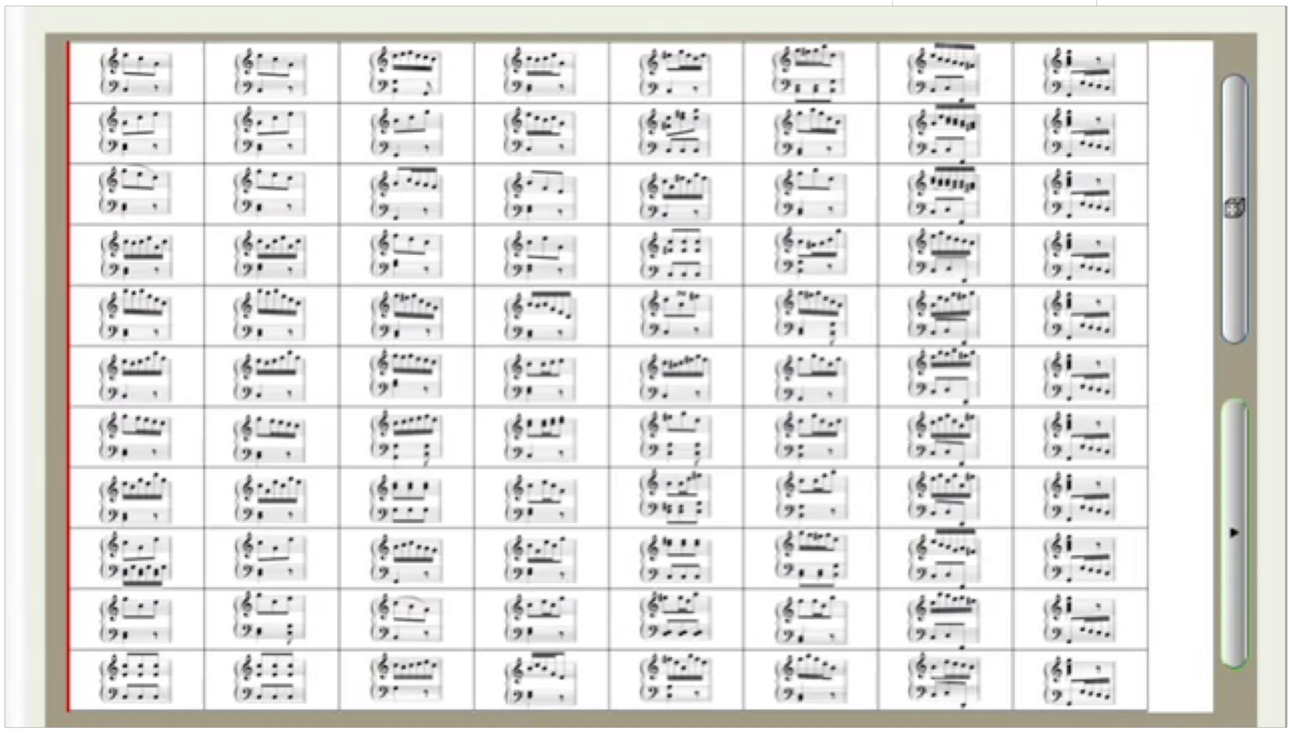
Algorithmic Composition

- Started in the 50's
- ILLIAC Suite [Hiller & Isaacson, 1957]
- Stochastic (Markov) models
- Filtering Rules



Mozart Dice Music

- Actually one of the First Documented Stochastic Music (1792)
- By Mozart (?) – Muzikalisches Wurfelspiel (Dice Music)
- Fixed Style (Wienna 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



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
  - Generation (ILIAC Suite)
  - Accompaniment: Harmonization, Counterpoint…
- Markov Chains
  - Generation (Random Walk, Constrained)
  - Style Imitation
- Neural Models (Deep Learning)
  - Prediction/Classification
  - Style Imitation

(Simplistic) Example of Generative Grammar for Blues

8-Measures -> First-4-Measures Second-4-Measures

First-4-Measures -> Opening-Cadence Opening-Cadence

-> Opening-Cadence’ Opening-Cadence

Second-4-Measures -> Middle-Cadence Opening-Cadence

Opening-Cadence -> | I | I |

-> | I | V |

Opening-Cadence’ -> | I | III |

-> | I | IV |

Middle-Cadence -> | I | IV |

-> | I | V |

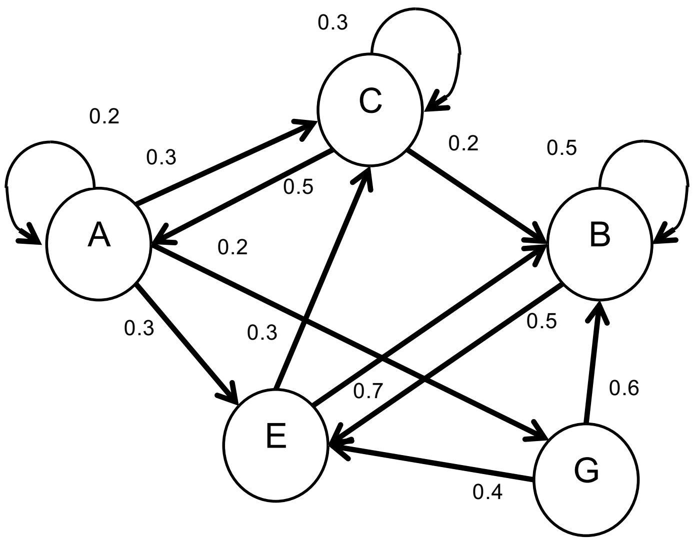
-> | IV | I | [John-Laird, 1991]

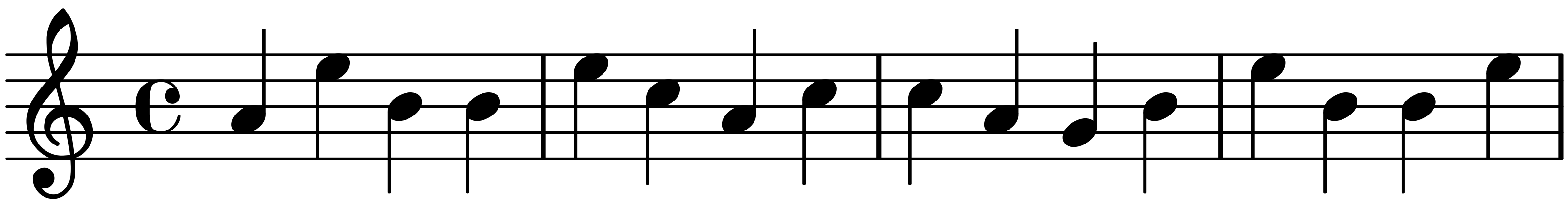
Ex:

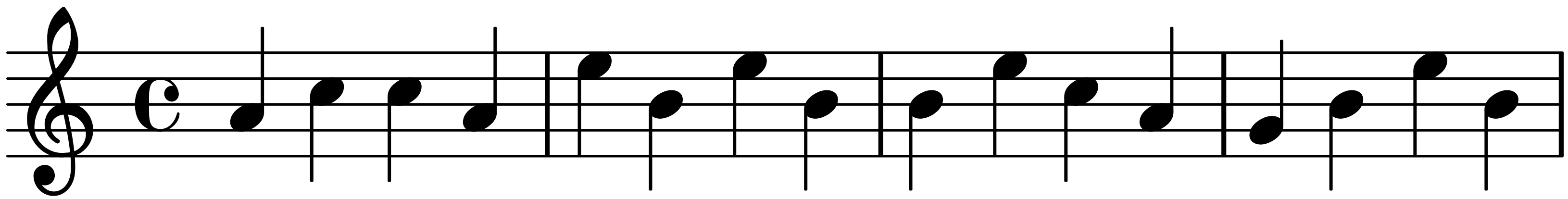
| I | I | I | I | IV | I | I | V |

| I | IV | I | V | IV | I | I | I |

(Simplistic) Example of Markov Model







Models

- Cellular Automata
  - Evolution
- Genetic Algorithms
  - Selection
- Case-based Reasoning
  - Similarity and Adaptation
- Planning
  - Path (Melody, Chord Sequence…) Construction
- …

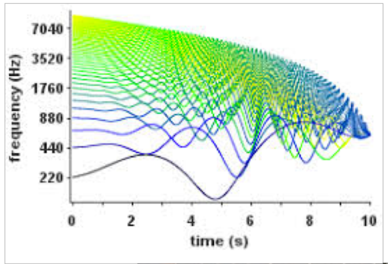
Importance of Randomness (Stochasticity)

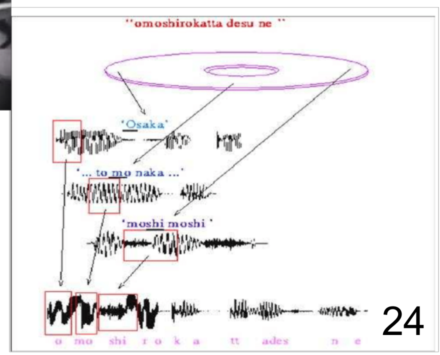
- Randomness/Stochasticity
  - Governed by Probabibilities
- To avoid Determinism
  - Otherwise Infinitely Repeats the Same Generation
- To be able to Generate Various Musical Pieces from a single Generative Model

- Ex: Mozart Dice-Music, Xenakis…
- Incorporated in most of Algorithmic Compositions

Also Models of Synthesis/Sound

- Additive
  - Sinusoids/Fourier^{-1}
- Substractive
  - Filtering
- Additive+Operators
  - Analog Synthesizers
- Physical
  - ex: Singing Voice, String…
- Frequency Modulation
  - Ex: Yamaha DX7
- Sampling
  - Initial Analog Version : Musique Concrète
- Granular
  - Arbitrary Objective
- Concatenative (Musaicing)
  - Source Objective
- …



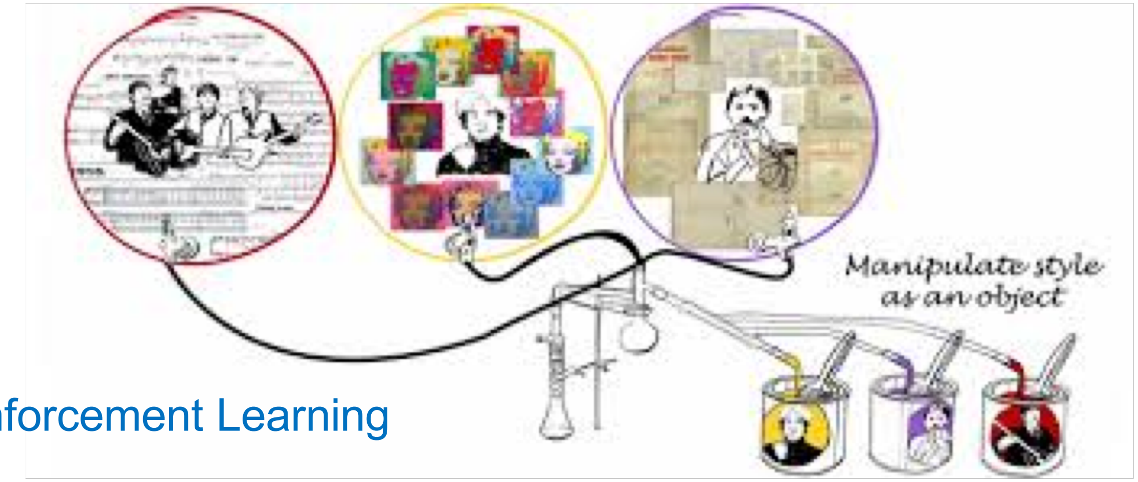


Handcrafted vs Learnt Models

- Handcrafted
  - Tedious
  - Error-Prone

- Automatically Learnt (Induction)
  - Markov models
  - Neural models

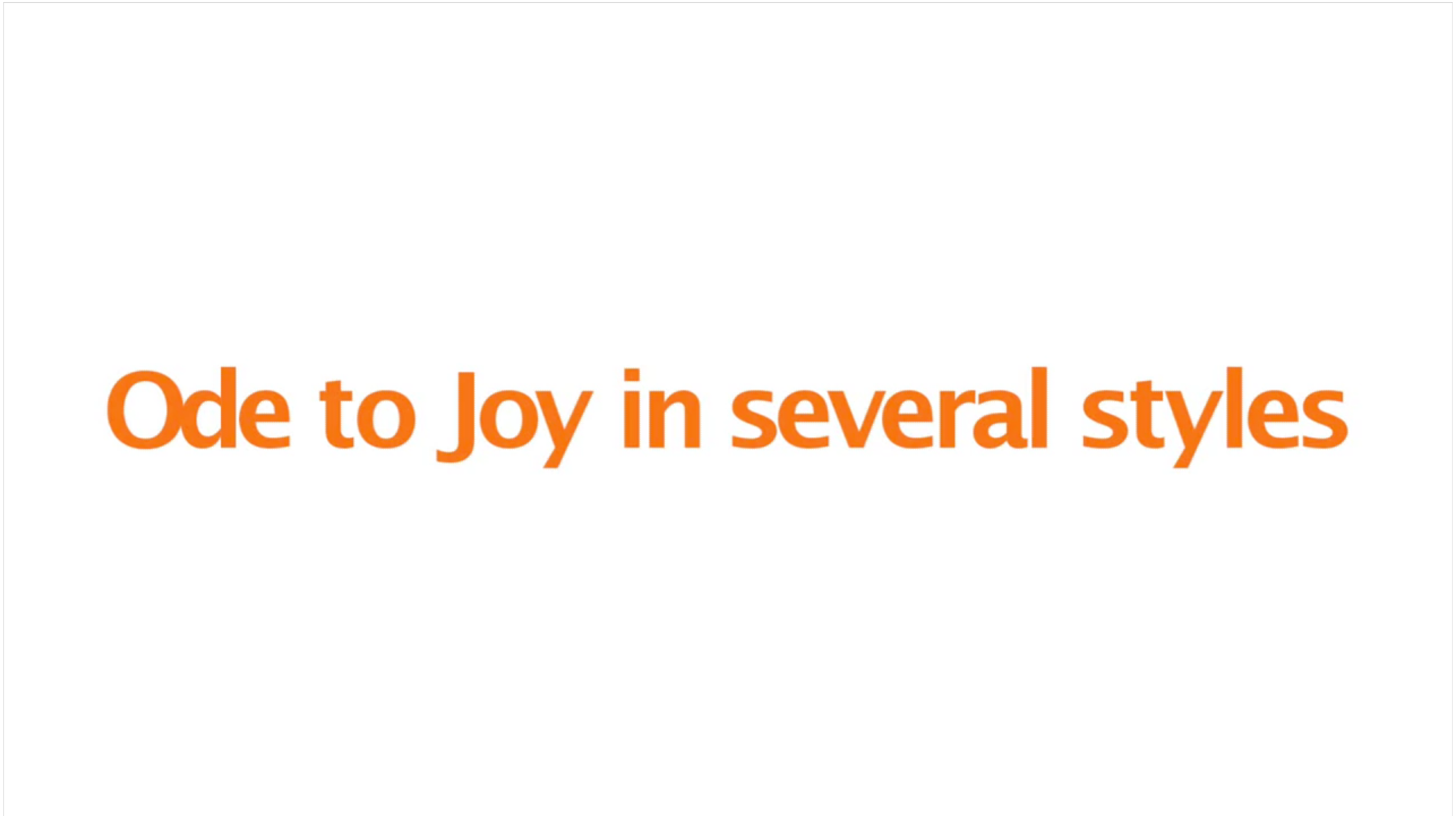
- Style Automatic Learned from a Corpus (Composer, Form, Genre…)



- Machine Learning Techniques
  - Neural Networks, Deep Learning, Reinforcement Learning
  - (and other models/techniques)

Flow Machines [Pachet et al. 2012]

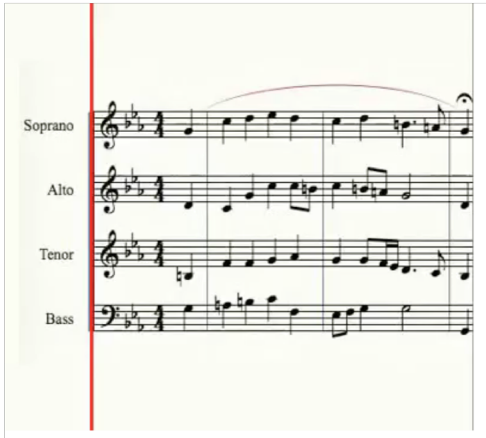
Reorchestration of Ode of Joy by DeepBach (and other techniques [Flow Machines])



Autonomous vs Assistance/Interactive Music Making

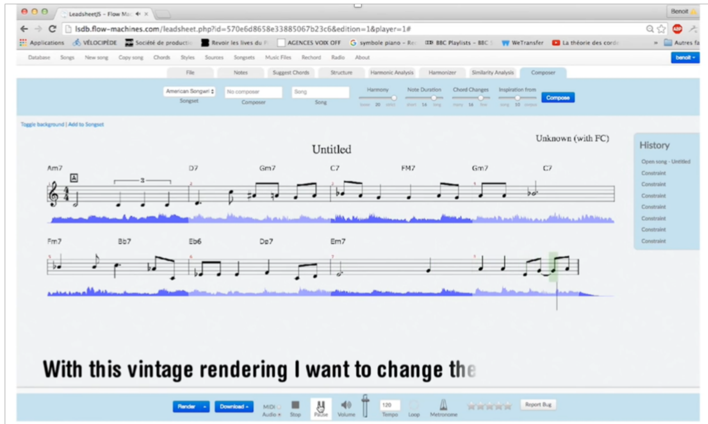
Autonomous vs Assistance Music Making

- Autonomous Generation/Interpretation
  - Turing Test
  - Ambient Music
  -



CompressorHeads (no AI Inside ☺)

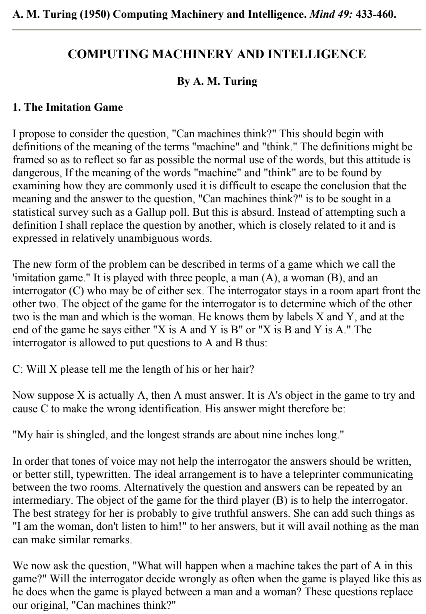
- Assistance to Human Composers and Musicians
  - Propose
  - Refine
  - Analyze
  - Harmonize
  - Produce
  - …



Bach Chorales Turing Test

- Autonomous Artificial Musicians

- Music Composition Turing test
  - Imitation Game Scenario [Turing, 1950]
  - Designed by A. Turing to explore the question "Can Machines think?"



A. M. Turing (1950) Computing Machinery and Intelligence. *Mind 49:* 433-460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game

I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous, If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game." It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart front the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either "X is A and Y is B" or "X is B and Y is A." The interrogator is allowed to put questions to A and B thus:

C: Will X please tell me the length of his or her hair?

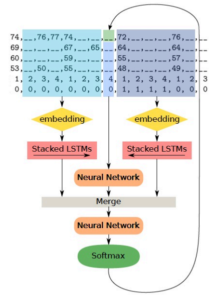
Now suppose X is actually A, then A must answer. It is A's object in the game to try and cause C to make the wrong identification. His answer might therefore be:

"My hair is shingled, and the longest strands are about nine inches long."

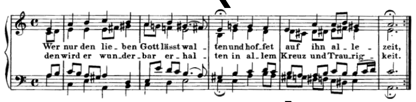
In order that tones of voice may not help the interrogator the answers should be written, or better still, typewritten. The ideal arrangement is to have a teleprinter communicating between the two rooms. Alternatively the question and answers can be repeated by an intermediary. The object of the game for the third player (B) is to help the interrogator. The best strategy for her is probably to give truthful answers. She can add such things as "I am the woman, don't listen to him!" to her answers, but it will avail nothing as the man can make similar remarks.

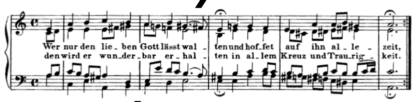
We now ask the question, "What will happen when a machine takes the part of A in this game?" Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, "Can machines think?"

(A) J. S. Bach



(B) DeepBach [Hadjeres et al., 2017]





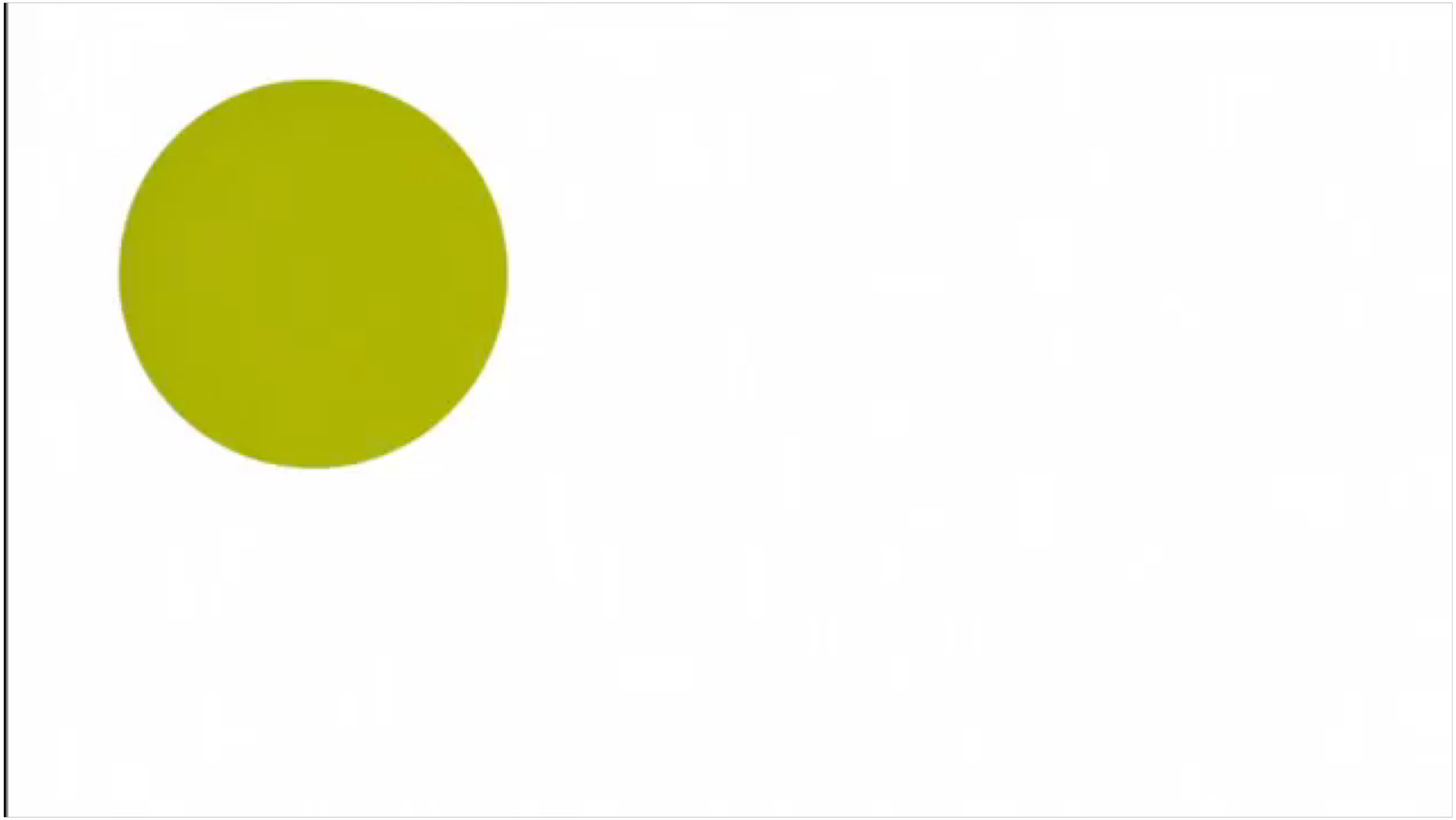
?

(C) Listener

- To evaluate artificial composers techniques
- To explore music cognition

Bach Chorales Turing Test

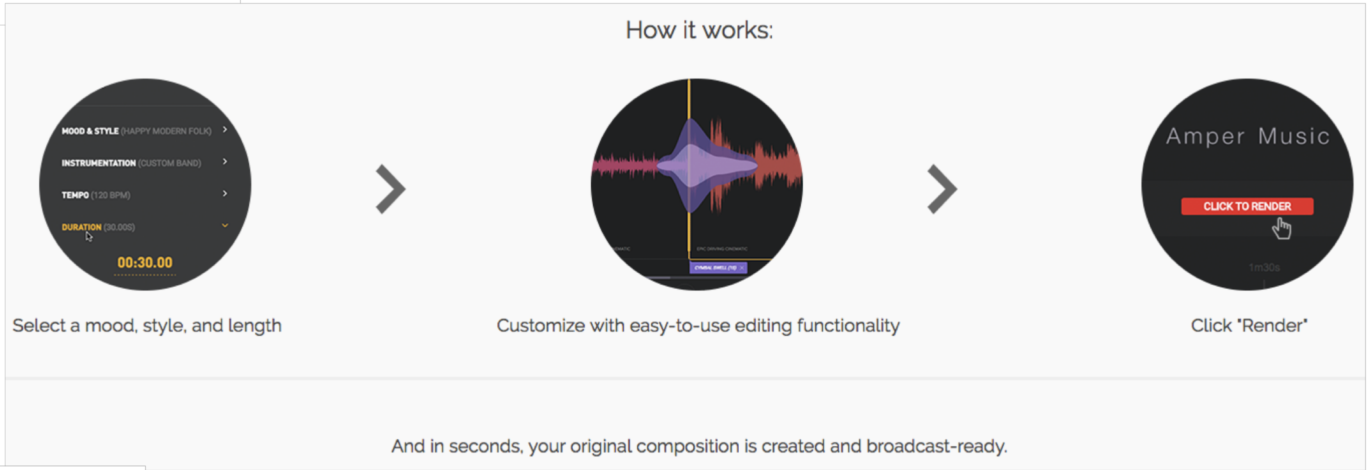
- February 2017, Dutch TV Channel
- Bach vs DeepBach Turing Test

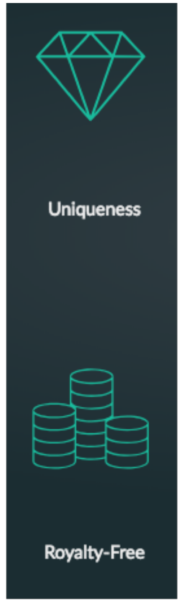


Autonomous Music Making

- Symbolic or/and Audio Music Generation

AMPER MUSIC





Jukedeck

Aiva Technologies

melomics media

  - https://www.youtube.com/watch?time_continue=11&v=aUFMFSI5qM8

Autonomous Music Making

- Symbolic or/and Audio Music Generation

- For Commercials and Documentaries
- Create Royalty-free or Copyright-buyable Music
- Based on Deep learning + Samples + Sound processing techniques

\+ Business model

-- Musical model

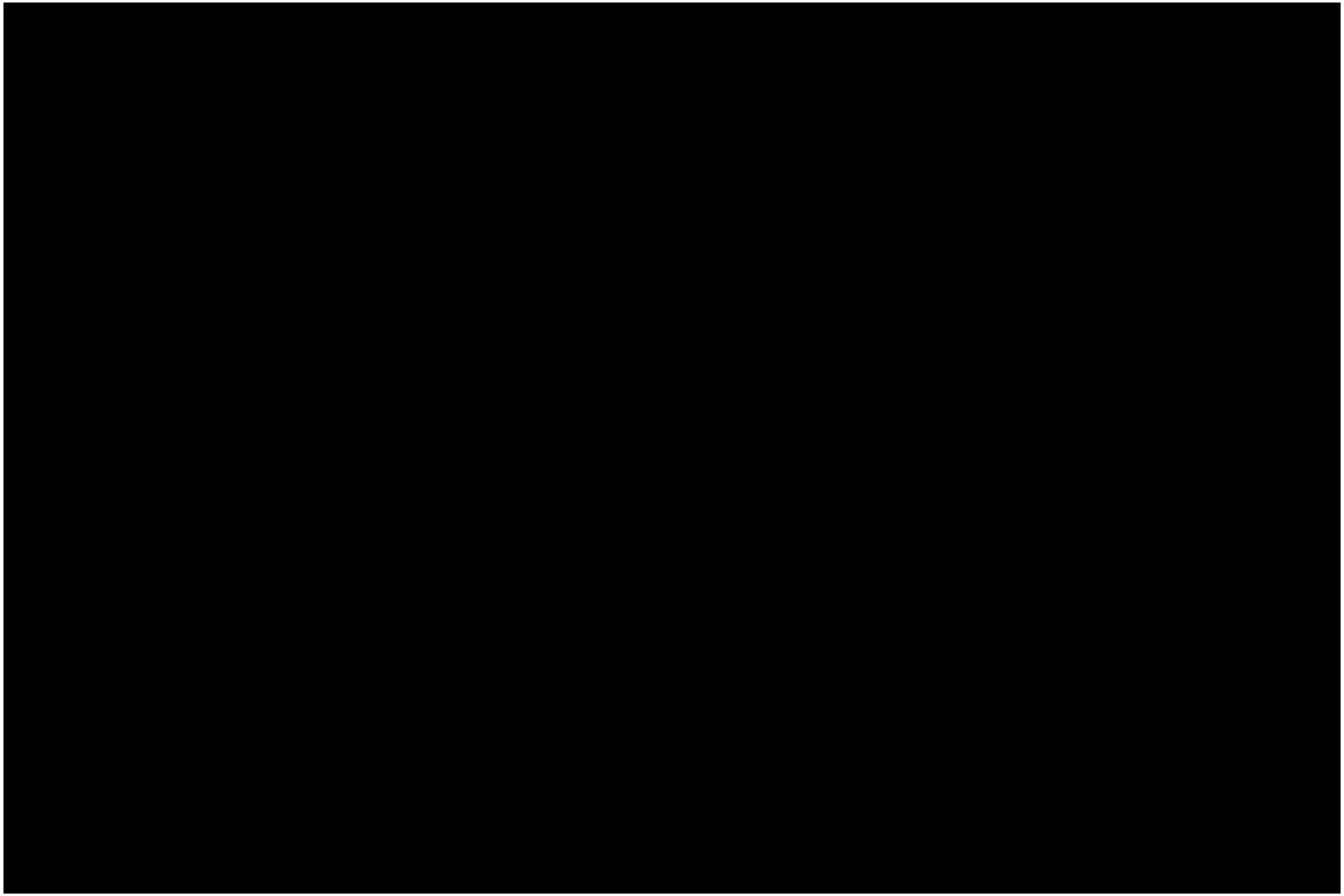
Interactive/Assistance

- Assistance to Human Composers and Musicians
  - Propose
  - Complete
  - Refine
  - Analyze
  - Harmonize
  - Adapt
  - Produce
  - Accompany
  - …

Objective and Evaluation [Pachet, 2019]

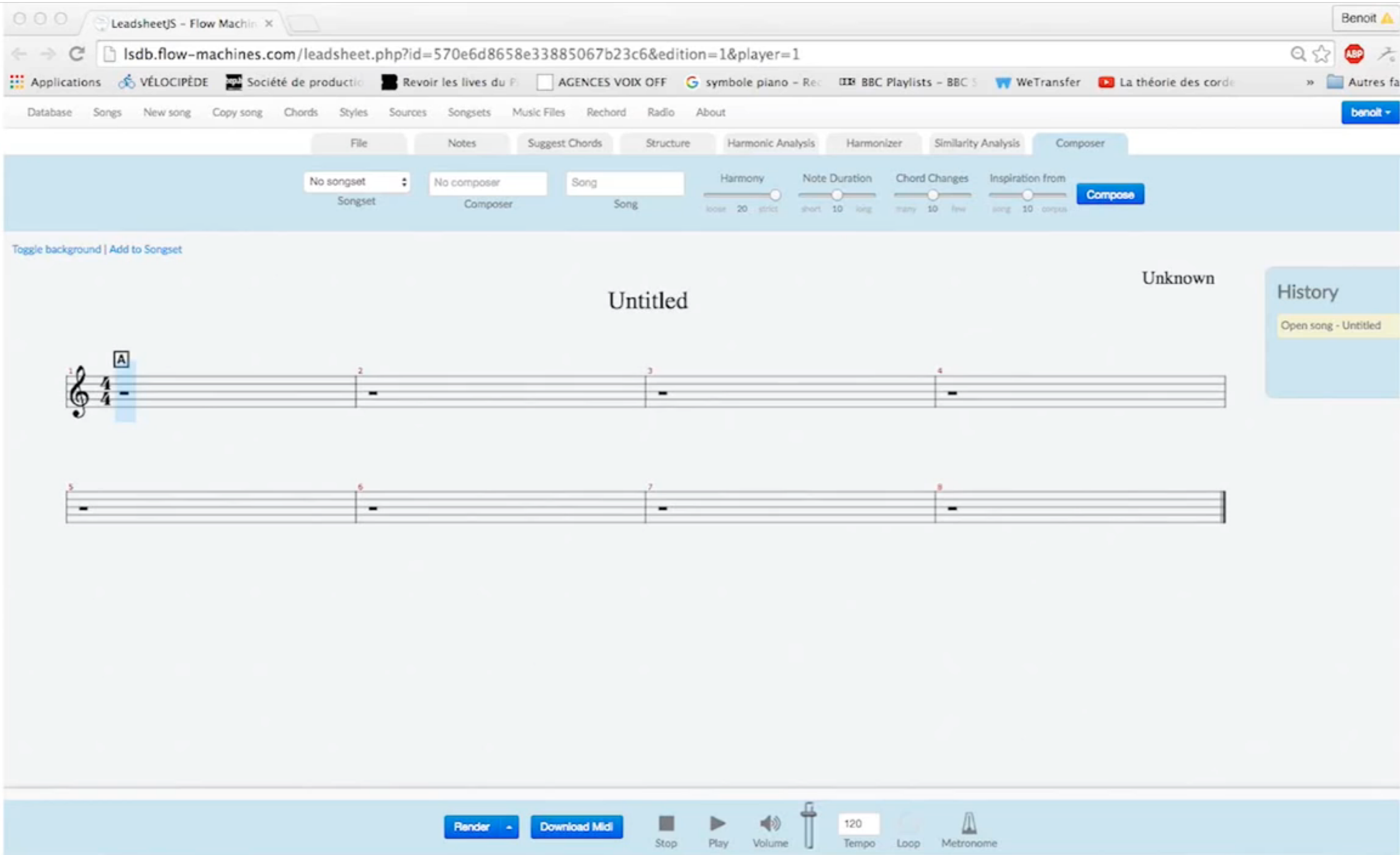
	Current Systems Autonomous Generalization-based	**Future** Systems **Augmentation/Assistance** **Creative**-incentived
Objective	Create music	Create music **not possible otherwise**
Evaluation	Please the listener	Please the **composer**
Risk	Conventional	**Surprising** But meaningful

Continuator [Pachet, 2002]



Pianist: Albert Veenendaal

FlowComposer [Pachet et al., 2014] – Demo (B. Carré)

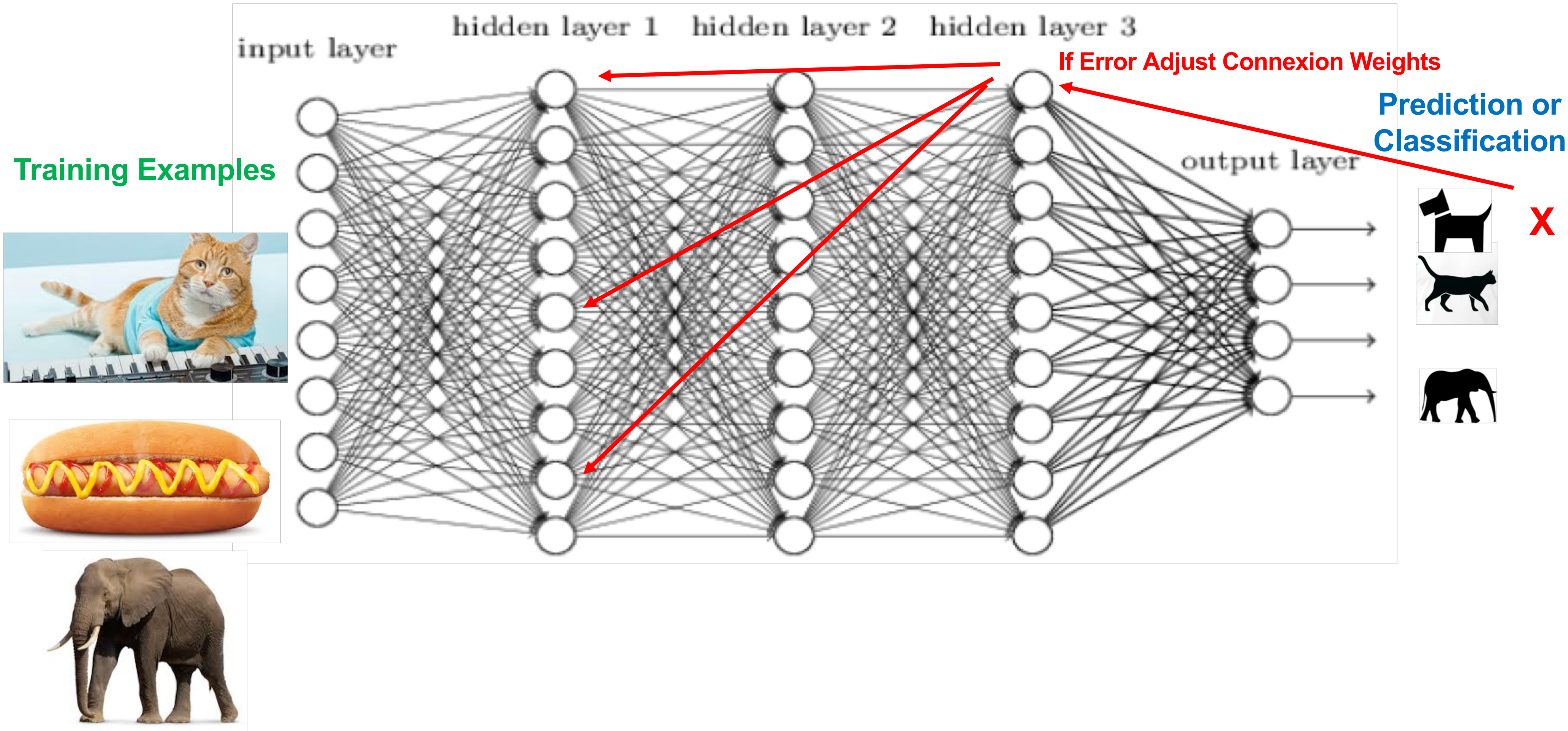


From Perceptron to Deep Learning

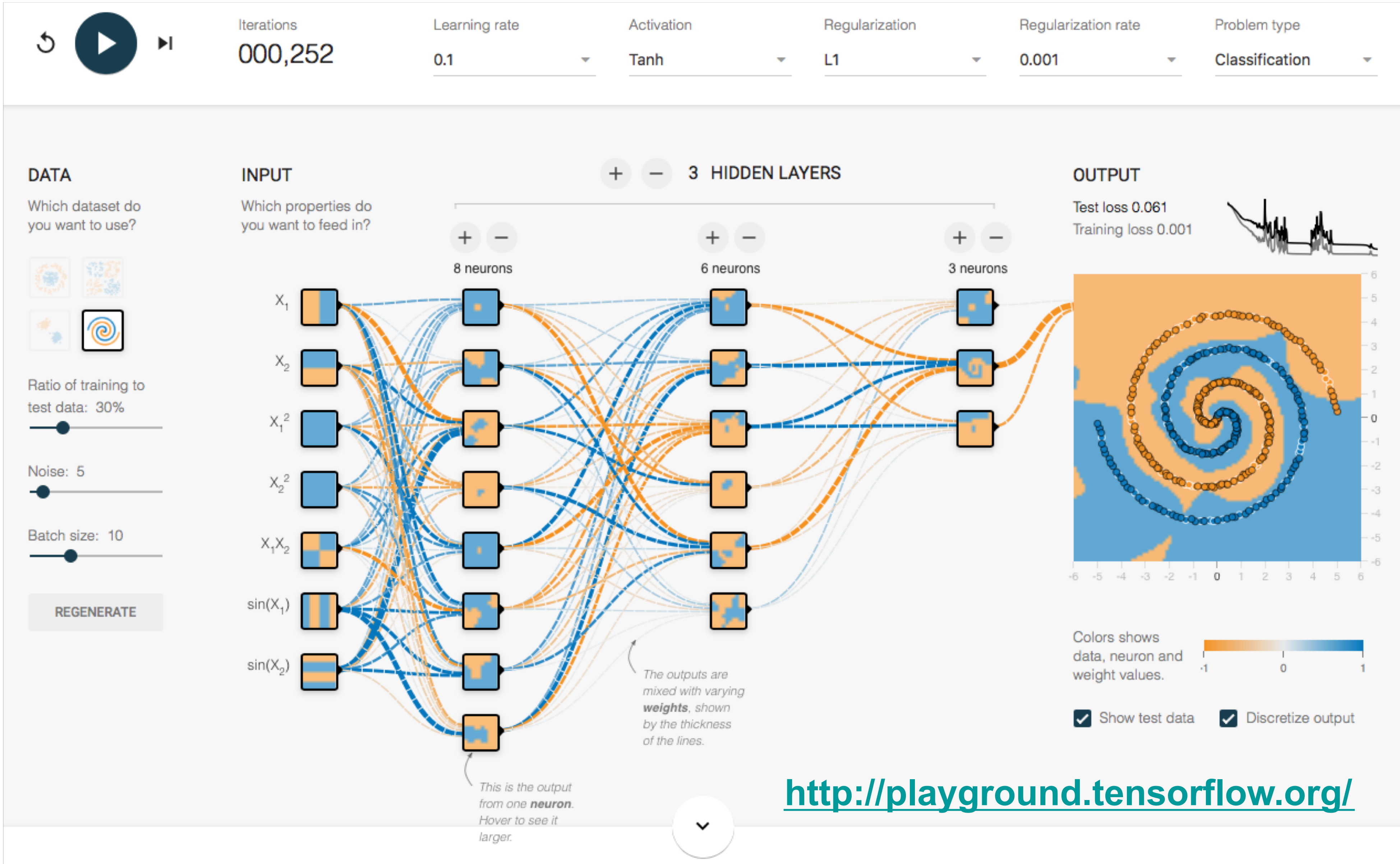
Linear Regression and Neural Networks

Neural Networks in One Slide

Principle – Error Prediction/Classification Feedback



Example (TensorFlow PlayGround)



http://playground.tensorflow.org/

Why Machine Learning?

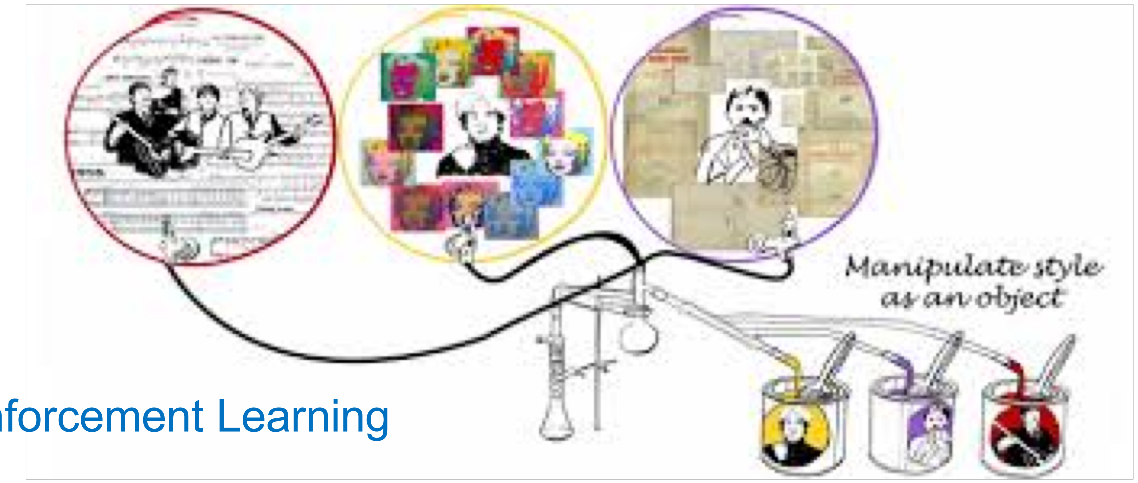
- Handcrafted
  - Tedious
  - Error-Prone

- Automatically Learnt (Induction)
  - Markov models
  - Neural models

- Style Automatic Learned from a Corpus (Composer, Form, Genre…)

- ++ Corpora Available on Line

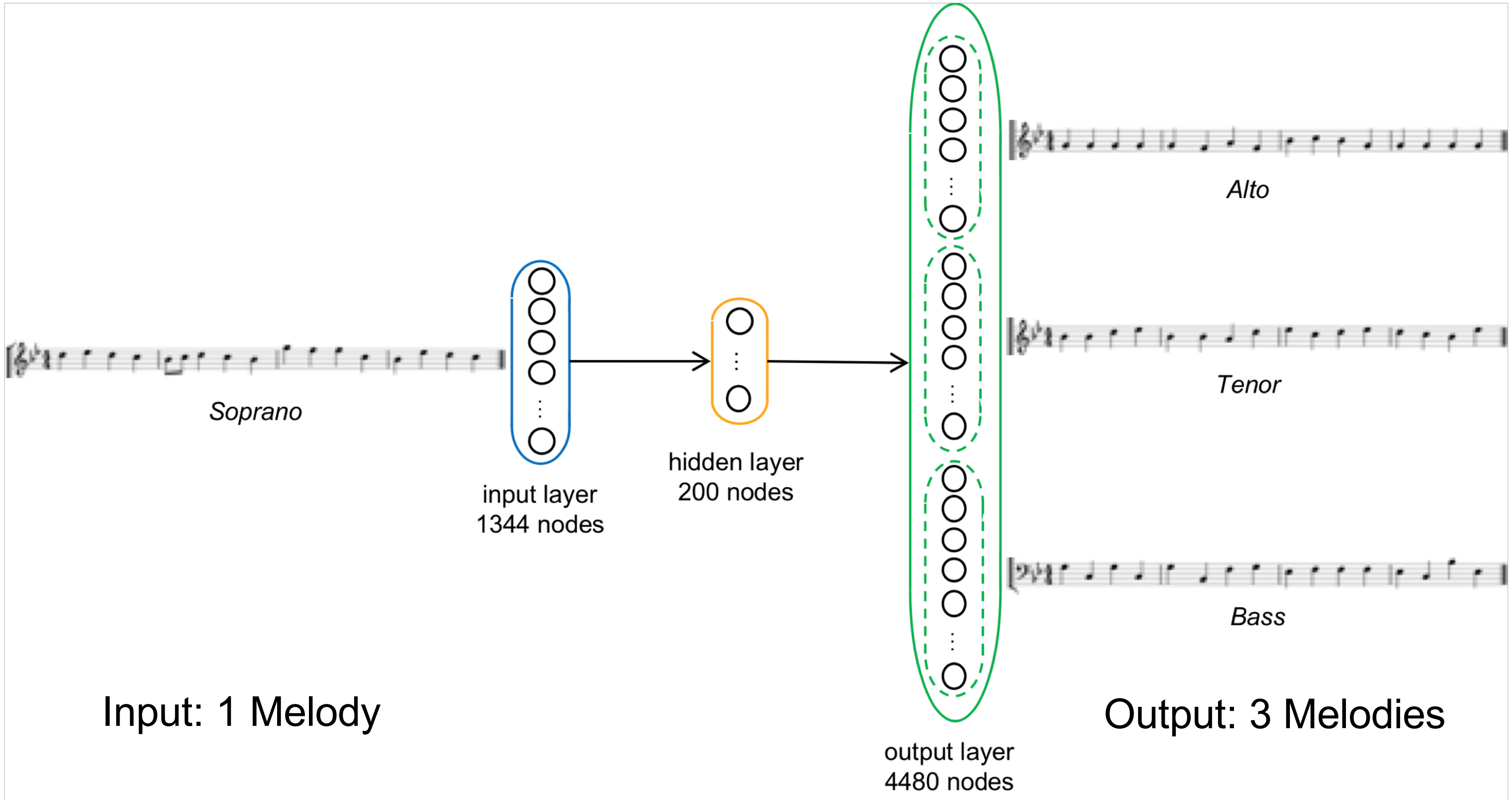
- Machine Learning Techniques
  - Neural Networks, Deep Learning, Reinforcement Learning
  - (and other models/techniques)



Flow Machines [Pachet et al. 2012]

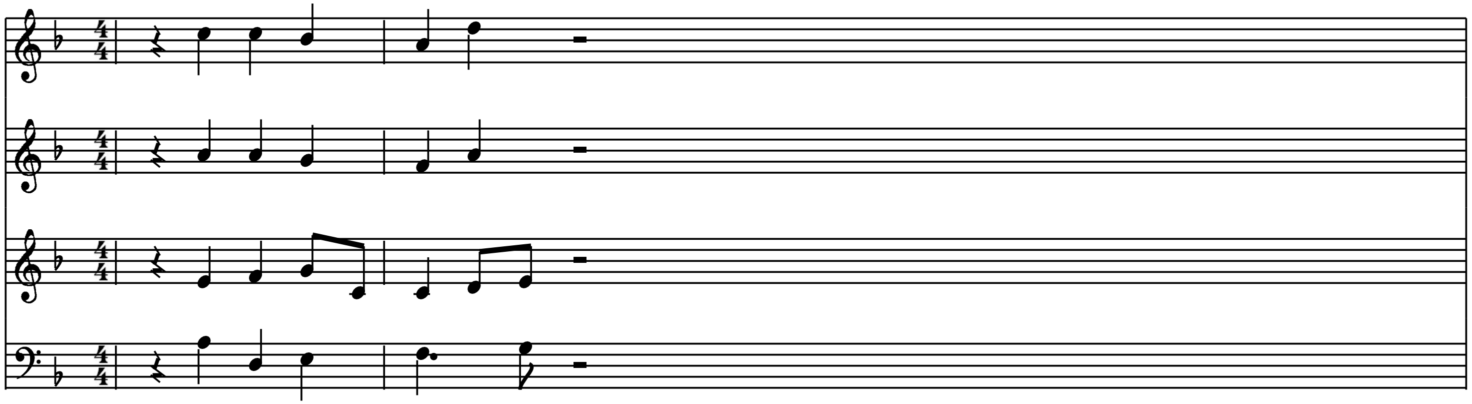
Direct Application – Ex: MiniBach [Hadjeres & Briot, 2017]

- Multilayer Architecture
- Classification Task (What Note?)
- Counterpoint (Chorale) generation
- Training on the Set of (389) J. S. Bach Chorales (Choral Gesang)

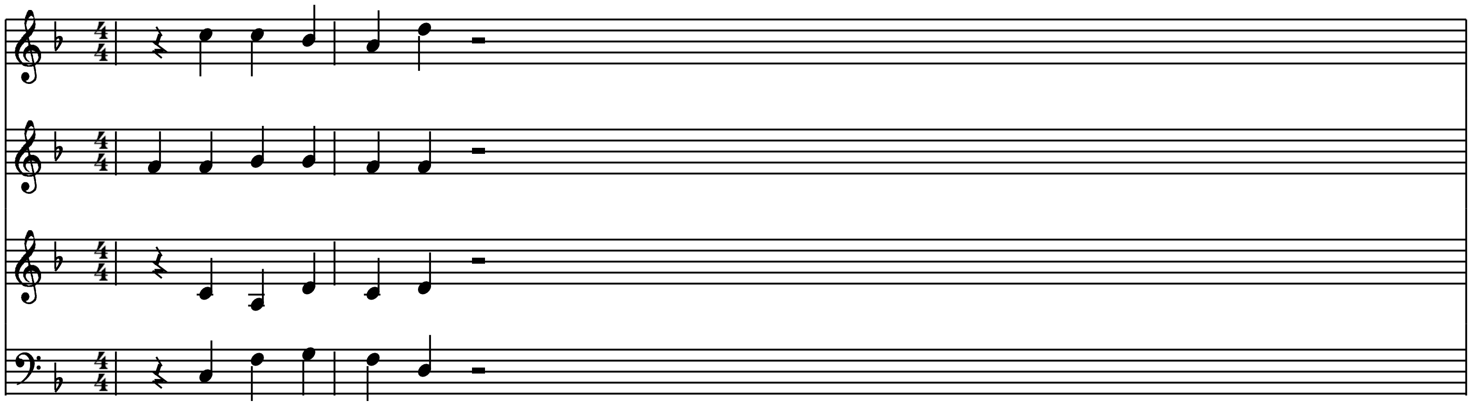


MiniBach Code & Examples

minibach-original-chorale



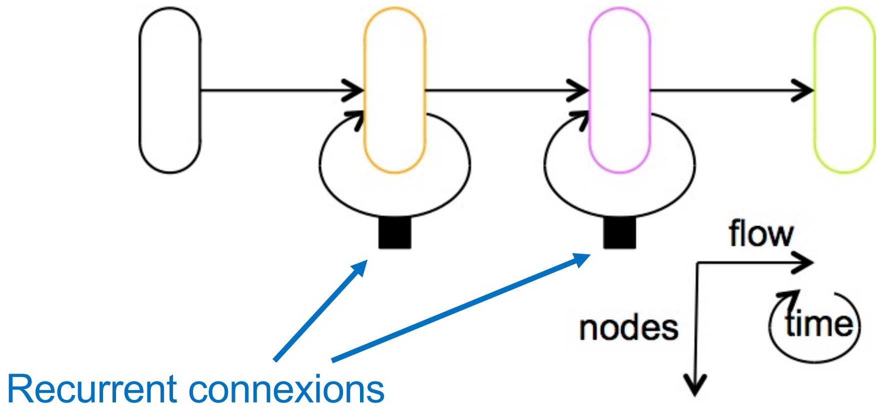
minibach-predicted-chorale

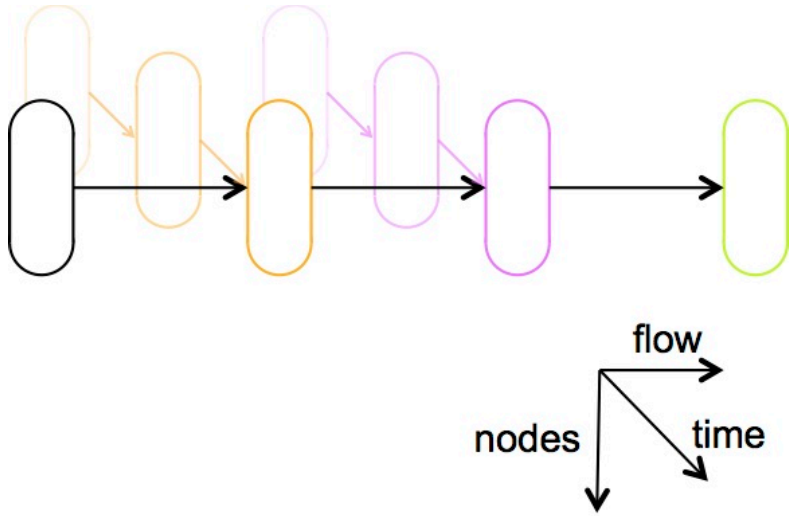


Reorchestration of God Save the Queen by DeepBach [Hadjeres, 2017]

Recurrent Network (RNN)

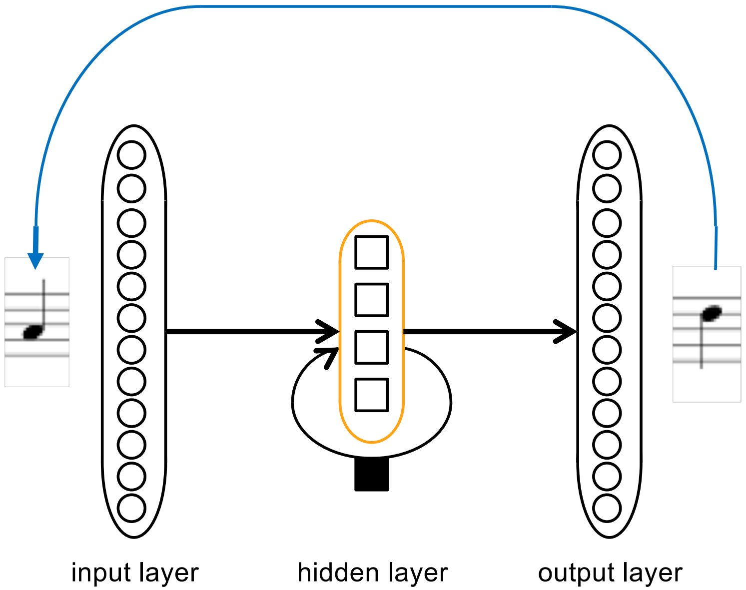
- Memorizes previous steps
- Can learn from previous step
- Predicts next step
- Can learn sequences

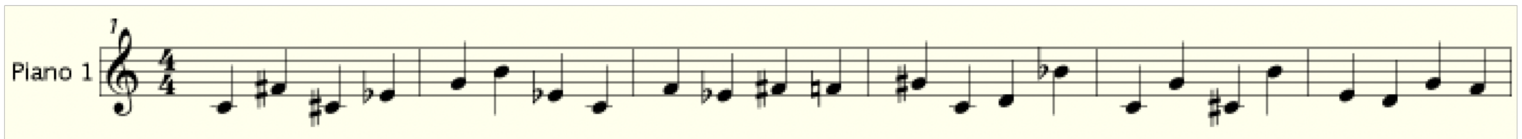




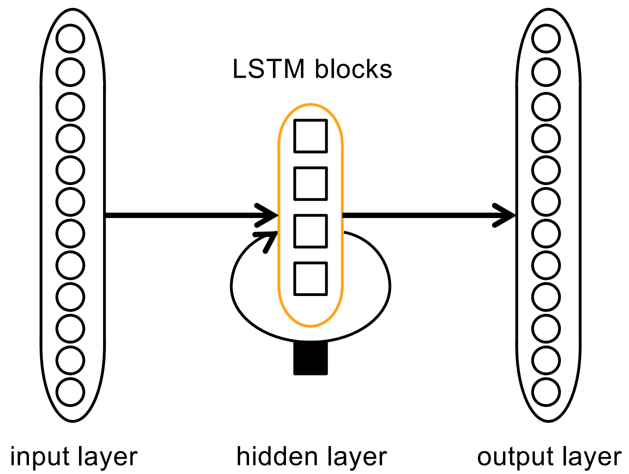
RNN Generation

- Iterated generation
  - Note by Note
  - Reinject Next Note to Produce Next Next Note, Etc.
  - Arbitrary Length



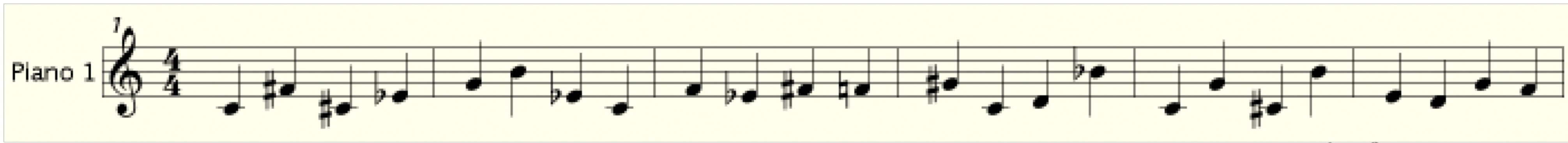


RNN – Iterative Feedforward – #1 Example



Synthetic corpus :
C major chord successive intervals
(3M, 3m, 3m, 2m)
4 examples only !

x *(note)*	*Interval*	y *(next note)*
C (1)	3M	E (3M)
E (3M)	3m	G (5)
G (5)	3m	B (7M)
B (7M)	2m	C (1)



RNN – Iterative Feedforward – #1 Example

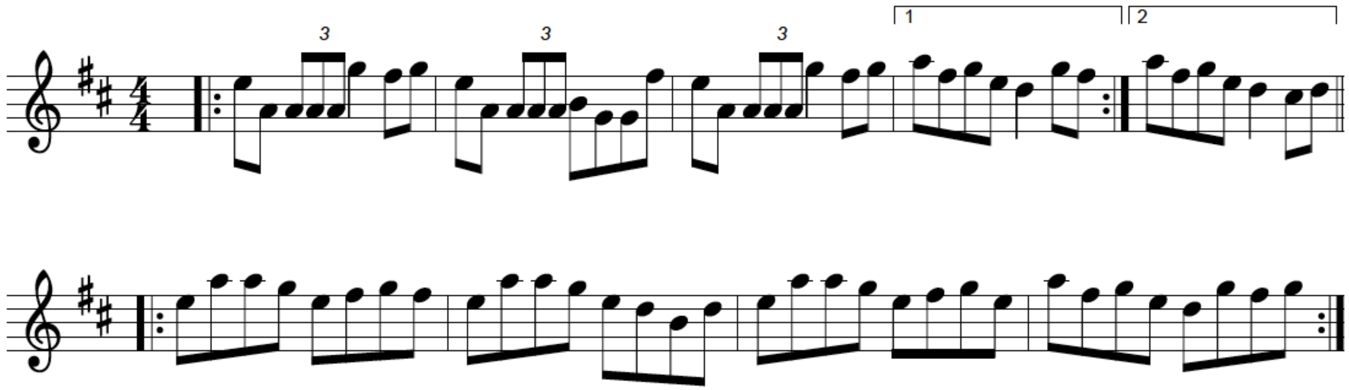
Keras Code

```
input_size = number_notes
lstm_layer_size = 32
output_size = number_notes

model = Sequential()
model.add(LSTM(lstm_layer_size,
               input_shape = (time_steps, input_size)))
model.add(Dense(output_size,
                activation = 'softmax'))

model.compile(loss = 'categorical_crossentropy')

model.fit(X_train,
          y_train)

melody_generated = []

for i in range(melody_length):
          note_one_hot = model.predict(np.array([[note_one_hot]]))[0]
          melody_generated.append(note_one_hot_to_name(note_one_hot))
```

RNN – Iterative Feedforward – #2 Example

- Ex: Celtic melody generation [Sturm et al., 2016]
- Celtic Folk Music Corpus (Melodies)
- Text Encoding (ABC Notation)

X: 1
T: A Cup Of Tea
R: reel
M: 4/4
L: 1/8
K: Amix
|:eA (3AAA g2 fg|eA (3AAA BGGf|eA (3AAA g2
fg|1afge d2 gf:|2afge d2 cd||
|:eaag efgf|eaag edBd|eaag efge|afge dgfg:|



RNN Celtic Melody Generation

- Iterated generation
  - Note by Note
  - Arbitrary Length
- Ex: Celtic melody generation [Sturm et al., 2016]
- Celtic Folk Music Corpus (Melodies)
- Text Encoding (ABC Notation)
- Ex. of Melody Generated

Played by a human accordeonist

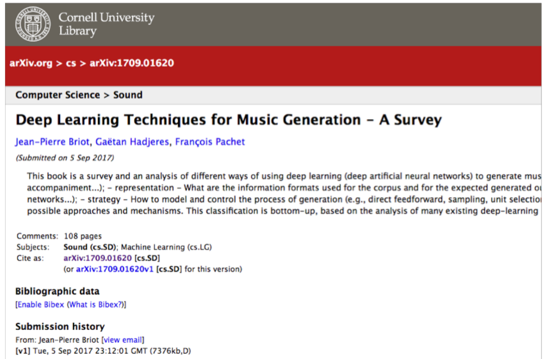
Conclusion/Prospects

- Deep Learning Progresses Fast
- Markov Models (and other Models) still Interesting
- Human Interaction – Assistance – Incrementality
- Creativity Incentive
- Plagiat

References

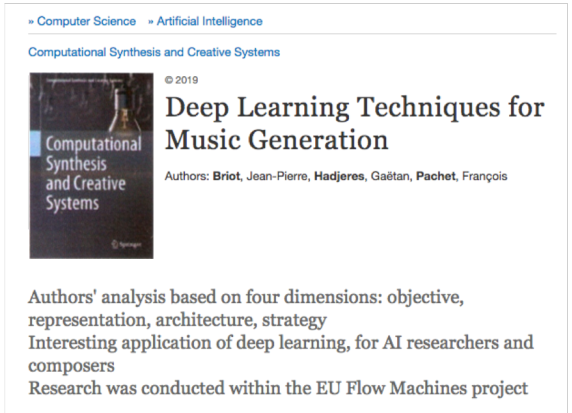
J.-P. Briot, G. Hadjeres, F.-D. Pachet, Deep Learning Techniques for Music Generation – A Survey, ArXiv:1709.01620, September 2017, November 2018, March 2019.

https://arxiv.org/abs/1709.01620



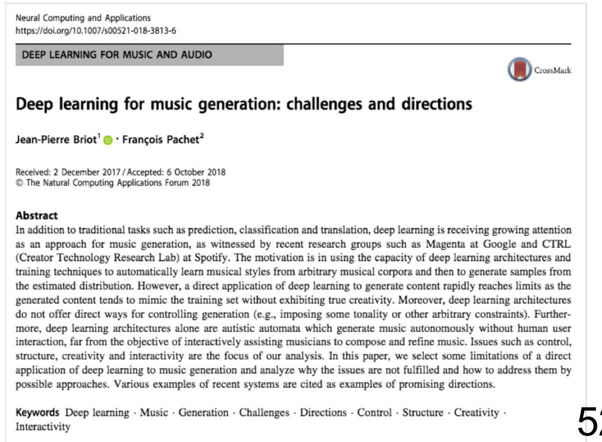
J.-P. Briot, G. Hadjeres, F.-D. Pachet, Deep Learning Techniques for Music Generation, Computational Synthesis and Creative Systems Series, Springer, 2019.

https://www.springer.com/br/book/9783319701622



J.-P. Briot, F. Pachet, Music Generation by Deep Learning – Challenges and Directions, Neural Computing and Applications (NCAA), Special Issue on Deep Learning for Music and Audio, October 2018.

https://doi.org/10.1007/s00521-018-3813-6



Acknowledgements

- Gaëtan Hadjeres
- François-David Pachet
- CNRS
- LIP6
- ERC Flow Machines
- Sony CSL-Paris
- Spotify CTRL
- PUC-Rio
- CAPES
- UNIRIO

Thank You – Questions