ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

ΦΙΛΟΣΟΦΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΟΥΣΙΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Artificial Intelligence Music Generators in Real Time Jazz improvisation: a performer's view

Δημήτριος Ε. Σμαΐλης

Επιβλέπουσα: **Αναστασία Γεωργάκη,** Αναπληρώτρια Καθηγήτρια
Δημήτρης Βασιλάκης, Μουσικός, Διδάσκων στο ΠΜΣ

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2020

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Artificial Intelligence Music Generators in Real Time Jazz Improvisation: a performer's view

Δημήτριος Ε. Σμαΐλης
Α.Μ.: 18209

Τριμελής Επιτροπή: **Αναστασία, Γεωργάκη,** Αναπληρώτρια Καθηγήτρια
Χριστίνα Αναγνωστοπούλου, Αναπληρώτρια Καθηγήτρια
Αρετή Ανδρεοπούλου, Επίκουρη Καθηγήτρια

Σημείωμα του συγγραφέα

Το δοκίμιο αυτό αποτελεί διπλωματική εργασία η οποία συντάχθηκε για το Τμήμα Μουσικών Σπουδών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και υποβλήθηκε προς εξέταση τον Φεβρουάριο του 2020. Ο συγγραφέας βεβαιώνει ότι το περιεχόμενο του παρόντος έργου είναι αποτέλεσμα προσωπικής εργασίας και ότι έχει γίνει η κατάλληλη αναφορά στην εργασία τρίτων, όπου κάτι τέτοιο ήταν απαραίτητο, σύμφωνα με τους κανόνες της ακαδημαϊκής δεοντολογίας.

Οι απόψεις που παρουσιάζονται στην παρούσα εργασία εκφράζουν αποκλειστικά τον συγγραφέα και όχι την επιβλέπουσα Καθηγήτρια.

Table of Contents

TABLE OF FIGURES 5

ACKNOWLEDGMENTS 7

ABSTRACT 8

INTRODUCTION 9

1. STATE OF THE ART: ALGORITHMIC MUSIC SYSTEMS FROM TETRACTYS TO OID. 11

1.1. Algorithmic music in the core of the history. 11

1.2. The use of computers in music generation. 12

1.3. Artificial intelligence and music generation 12

2. AI MUSIC GENERATORS: OMAX-IMPROTEK-DJAZZ 14

2.1. OMax: towards experimental music 14
2.1.1. Factor Oracle Basics 15
2.1.2. OMax in the context of live performance 17

2.2. ImproteK: introducing structured improvisation 19
2.2.1. Scenario & Memory 20
2.2.2. Navigation and Synchronization 21
2.2.3. Online and offline learning 22
2.2.4. ImproteK at work 22

2.3. Djazz: towards a consumer product 24

3. REFLECTING ON THE MUSICALITY OF OID COMPUTER MUSIC SYSTEM: A CRITICAL APPROACH 26

3.1. Performance outcomes of OID and the Elements of Music 29
3.1.1. Beat, Meter, Rhythm, Groove, Silence 29
3.1.2. Harmony 32
3.1.3. Melody 34
3.1.4. Timbre 37
3.1.5. Dynamics 37
3.1.6. Texture 38
3.1.7. Structure – Form 38

4. DISCUSSION 42

5. CONCLUSION 44

6. PERFORMANCE ANALYSIS 45

6.1. Discussion **45**

6.2. Chameleon **46**
6.2.1. Musical notes 46
6.2.2. Background 47
6.2.3. Technical notes 47
6.2.4. Performance structure 47

6.3. OMax & Saxophone Duet **48**
6.3.1. Technical notes 48
6.3.2. Performance structure 48

REFERENCES **51**

Table of Figures

FIGURE 1: THE FACTOR ORACLE FOR STRING ABBBAAB. SUFFIX LINKS ARE IN DOTTED LINES (ASSAYAG & BLOCH, 2007). 15
FIGURE 2: THE FACTOR ORACLE. A GRAPH BASED ON SYMBOLIC SEQUENCES. THE DASHED BACKWARD ARROWS POINT AND CONNECT THE SIMILAR PATTERNS. 16
FIGURE 3: SUFFIX LINKS AND LRS. A SEARCH MAP OF THE SEQUENCE ABOVE AND HOW THE PATTERNS CAN CONNECT TO EACH OTHER, WHICH ARE SIMILAR TO WHICH AND WHERE THEY ARE IN THE TIMELINE. 16
FIGURE 4: SUFFIX LINKS TREES. ALL THE PATTERNS CONNECTED IN THIS THESE TREES HAVE THE SAME SUFFIX, THE SAME ENDING. THESE ARE THEN USED TO GENERATE VARIATIONS. THE SUFFIX TREE GUARANTEES THAT THE CONCATENATION WILL BE SMOOTH (LEVY, 2013). 16
FIGURE 5: THE SYSTEM READS THE SEGMENTED DATA. 17
FIGURE 6: WHEN IT IS TIME TO MAKE A VARIATION OF THE LEARNT MATERIAL AN ANTICIPATION WINDOW EXPLORES ALL GRAPHS AND TREES IN REGARDS TO THE DESCRIPTORS (PITCH, TIMBRE, HARMONY) 17
FIGURE 7: DESCRIPTORS CONSIDERED 17
FIGURE 8: ALL SOLUTIONS ARE GATHERED AND WEIGHTED, THEN THE SYSTEM JUMPS TO A SIMILAR PATTERN AND CONTINUOUS ON. 17
FIGURE 9: LEARNING THE *LIVE ORACLE* 20
FIGURE 10: USING THE SCENARIO TO INTRODUCE ANTICIPATION IN THE MUSIC GENERATION PROCESS. 23
FIGURE 11: TABLE OF FUNCTIONS OF OMAX, IMPROTEK AND DJAZZ 25
FIGURE 12: OCTAVE DISPLACEMENT AND CHORD TONE PLACEMENT ON BEATS 1 AND 3, BERT LIGON ON 'CONNECTING CHORDS WITH LINEAR HARMONY' 30
FIGURE 13: FIRST FOUR BARS FROM THE BRIDGE OF AN IMPROVISATION ON *AN OSCAR FOR TREADWELL* (RHYTHM CHANGES FORM). THE 'IMPROVISATION' LINES SHOWS THE CHORD AND DEGREES PLAYED IN THE GENERATED IMPROVISATION; THE 'MEMORY' LINES SHOWS THE CHORD AND DEGREES ON WHICH THE DIFFERENT PARTS OF THE MELODY USED TO GENERATE THE IMPROVISATION WERE ACTUALLY LEARNT. WE SEE THAT WHEN GENERATING IMPROVISATIONS, WE CAN REACH PLACES IN THE MEMORY WITH A DIFFERENT CHORD LABEL. THE MELODY STILL MAKES SENSE IN ITS CONTINUITY AND HARMONICALLY BECAUSE THE MULTI-LEVEL LABELS FROM THE MEMORY AND FROM THE MULTI-LEVEL PROGRESSIONS ARE EQUIVALENT. THE GENERATED IMPROVISATIONS ARE THEREFORE ENRICHED WITH A NEW FORM OF CREATIVITY (DEGUERNEL ET AL, 2017). 33
FIGURE 14: 7TH TO THE 3RD, CHARLIE PARKER ON 'DONNA LEE' 35

FIGURE 15: CHROMATIC ENCLOSURE, MICHAEL BRECKER ON 'SOFTLY AS IN A MORNING SUNRISE' 35
FIGURE 16: USE OF BEBOP SCALE, CHARLIE PARKER ON 'DONNA LEE' 35
FIGURE 17: RESOLVING BY STEP, JOHN COLTRANE ON 'LAZY BIRD' 35
FIGURE 18: ASCENDING CONTOUR, MICHAEL BRECKER ON 'SOFTLY AS IN A MORNING SUNRISE' 35
FIGURE 19: RESOLVING ON THE 6TH ON THE OFF-BEAT OF TWO, CHARLIE PARKER ON 'DONNA LEE' 35
FIGURE 20: BASS AND DRUMS ON THE FIRST SECTION OF "CHAMELEON" 46
FIGURE 21: GRAPHIC SCORE BY CARL BERGSTROM-NEILSEN 50

Acknowledgments

I would like to thank my supervisor Anastasia Georgaki for her valuable guidance. I particularly appreciated her versatile view, her wide knowledge of current trends, her experience in academic writing, the organization of relative events and conferences and the invitation of key speakers during the course. I would also like to thank my university music teachers Dimitri Vassilakis for the useful conversations and the implementation of performances using Artificial Intelligence generators, Antonis Ladopoulos for his motivation, efficiency and teaching professionalism and Areti Andreopoulou for her guidance on technical issues and her methodical, systematic and organized lessons. Thanks also goes to, my fellow students from Kapodistrian University of Athens Nastazia Beikof, Natalia Kotsani and Tilemachos Moussas, the IRCAM team, that is, Jerome Nika, Benjamin Levy and Raphael Imbert all of whom I had the opportunity to meet, learn from and perform with using Artificial Intelligence music systems.

Abstract

A highly controversial entrance of Artificial Intelligence (AI) music generators in the world of music composition and performance is currently advancing. A fruitful research from Music Information Retrieval, Neural Networks and Deep Learning, among other areas, are shaping this future. Embodied and non-embodied AI systems have stepped into the world of jazz in order to co-create idiomatic music improvisations. But how musical these improvisations are? This dissertation looks at the resulted melodic improvisations produced by OMax, ImproteK and Djazz (OID) AI generators through the lens of the elements of music and it does so from a performer's point of view. The analysis is based mainly on the evaluation of already published results as well as on a case study I carried out during the completion of this essay which includes performance, listening and evaluation of generated improvisations of OMax. The essay also reflects upon philosophical issues, cognitive foundations of emotion and meaning and provides a comprehensive analysis of the functionality of OID.

Introduction

Twentieth century technological inventions have transformed the production and distribution of music. Over the century, recording and audio processing have provided musicians with assistive tools that are still shaping the future of music pedagogy, performance, production, distribution and consumption. Twentieth first century further enhanced music by introducing Artificial Intelligence (AI) which, by many, is considered as one of the most important inventions in the history of humankind justifiably claiming a place alongside electricity, steam engines and the internet. In terms of composition and performance the focus of this new AI music-technological wonder is on developing creative interactive systems for both, hence shifting music technology's status from assistive to co-creative. Music Information Research (MIR) is at the forefront of these developments.

Until recently, interactive music systems have been used in the performance of experimental music where there are few traditions or constraints. Concerns of synchronization, harmonic structure, adherence to predetermined forms didn't have to be encompassed. Inevitably, musicians experimenting with AI systems have asked for more in terms of musical abilities. Hence, recent developments consider idiomatic performance, in the sense of both, composition i.e. Flow Machines and improvisation i.e. OID, rhythmic performance, synchronization and more.

This essay reflects on the use of OMax-Improtek-Djazz system, from now on referred as OID, from a human jazz musician and performer's point of view in regards to its melodic improvisation capabilities and musicality. OID is a new technology which allows humans and machines to improvise real-time music side by side. It uses Factor Oracles which is classified as the most successful technique for machine improvisation (Siphocly et al, 2019). To my knowledge, the system is one of a few that handles audio, with most systems i.e. Google's Magenta using MIDI. OID is supported by a human who controls its parameters and guides the system. My focus is on the responsiveness of OID in regards to the elements of music in real-time performance. Hopefully, this essay will contribute further to the constructive musicological dialogue around music generators.

The location, classification of the several versions of OID has been done through extensive research which included but not limited to: published scientific papers, forums, participation to conferences, related seminars and discussions with practitioners. The system was tested using OMax4x as, to my knowledge, this is the only software currently available and functional with my set-up.

In the first section, I consider the state of the art of Algorithmic music generation. The section is divided in three categories spanning from ancient Greek music to current Artificial Intelligence (AI) technologies, providing a short but solid overview of Algorithmic music. In the second section, in order to gain an understanding of OID's music generation processes, I examine the development of its several versions pointing on strengths and drawbacks. A basic comparative analysis of the several versions is taking place throughout the section. In the third section I reflect on OID from a cognitive point of view discussing concepts such as 'emotion', 'meaning', 'tension' and 'resolution'. I also make references to the ethos of machine music generation and to the debate between 'deep learning vs deep understanding'. Furthermore, even though the project is still under development, I analyse OID's generation capabilities in regards to the elements of music. It is my belief that the elements of music, apart from its pedagogical benefits, can provide a musicological evaluation terrain upon which a fruitful critic on any such system may develop. In the fourth section, I discus the matter in terms of musicality, cognition and research. In the fifth section, I finally conclude with a suggestion for further development. In the sixth section, I am providing analysis notes regarding my performance in the context of the Master's Programme 'Jazz Improvisation and Contemporary Practices' of the National and Kapodistrian University of Athens.

1. State of the Art: Algorithmic music systems from Tetractys to OID.

It is not the scope of this essay to provide a historical overview of algorithmic music, however, mainly for better comprehension of current systems, I cite examples of automated composition in the core of the history of Western music. For an overview of algorithmic music refer to Cope (2000), McLean and Dean (2018), Maurer (1999), Alpren (1995) and Burkholder et al. (2019). For computer music in general refer to Roads & Strawn (1996).

1.1. Algorithmic music in the core of the history.

Music and maths were always connected. Tetractys, a concept provided by Pythagoras demonstrates that western music is governed by maths through which the control of both inanimate and animate 'Cosmos' is possible. Pythagoras' 'Sectio Canonis' explaining the 'Monochord' and its fractions sufficiently demonstrates this assumption. Ancient Greeks based their musical system in formalisms 'algorithms'. Their music, although not algorithmically composed, showed a tendency of these formal extra-human processes. Greek Philosophy, spread to the Western world by philosophers such as Boethius, constituted the basis of Baroque music. Canonic composition of the late 15th century is an example of an early algorithmic composition (Burkholder et al, 2019). A challenging idea that was known and tried out before by composers such as Ph. Em. Bach and J. Haydn (Ruttkay, 1997), was Mozart's musical game 'Musikalisches Wurfelspiel' (Musical Dice Game), in which dice are used to select randomly from a number of possible arrangements of each bar in a Viennese minuet. 'Spiegelkanon' (mirror canon) by Mozart is also an example of this tendency. Later, in the post-world war II era, 12 tone and serial music tried to control all parameters of music and to objectify and abstract the compositional process as much as possible. John Cage, like Mozart, utilized randomness through algorithms in many of his compositions, i.e. *Reunion*, performed by playing chess on a photo-receptor equipped chessboard (Schwartz, 1993).

1.2. The use of computers in music generation.

The use of computers in algorithmic music composition came in the late 40's by Alan Turing himself who discovered that the different patterns emitted by his computer sound like different musical notes (Copeland & Long, 2017). In 1951, Christopher Strachey developed a program that generated Britain's national anthem "God Save the King" (Foy, 1974). Later, in 1955-56, Lejaren Hiller and Leonard Isaacson at the University of Illinois, using the Illiac, a high-speed digital computer, succeeded in programming basic material and stylistic parameters which resulted in the *Illiac Suite* in 1957. Since then, Yannis Xenakis and Karlheinz Stockhausen among others have been experimenting with algorithmic music composition (Alpern, 1995). Systems of music compositional techniques such as M by Joel Chadabe, GenJam by John A. Biles, Voyager by George Lewis, Ramses by Steve Coleman are also considered related to algorithmic music generation without, however, using machine learning schemes (Assayag et al, 2006).

1.3. Artificial intelligence and music generation

Nowadays, AI systems, apart from obeying to specific rules, have the capacity to create its own grammar and database of rules, in essence, a capacity to 'learn' either from a corpus or on-the-fly. An early such system is David Cope's 'Experiments in Musical Intelligence' (EMI). EMI is based on a large database of style descriptions which consists of rules of different compositional strategies and has been used to automatically compose music in the styles of Bach, Mozart, Bartók, Brahms, Joplin and many others (Cope, 1989).

In this new machine learning era there are several ongoing projects. IRCAM's OID is aimed at the generation of real-time improvisation by reinjecting captured audio or MIDI or through the use of corpus according to given descriptors. Google's Magenta uses deep learning models to generate melodies, rhythms and grooves (Roberts et al, 2019). Sony's Flow Machines is aimed at achieving augmented creativity of artists by co-composing music providing musicians with interactive computer programs that let users *play* with styles (Ghedini et al, 2016). Georgia tech's Robot named Shimon, a socially interactive and improvisational robotic marimba player can act within an interactive musical jam session among human and robotic musicians (Weinberg et al, 2009). For an overview of current systems one can see Tatar and Pasquier (2019), Gifford et al. (2018), Herremans et al. (2017), Kirke and Miranda (2013), Edwards (2011). For a survey on deep-learning techniques refer to Briot et al. (2017).

The study of co-creative human-computer music practice has been scientifically examined by several authors. Dannenberg (2012) for example has created an all-inclusive field, namely Human Computer Music Performance (HCMP), which concerns the study of music performance by live human performers and real-time computer-based performers. Among the goals of HCMP is the creation of:

> *"…a highly autonomous artificial performer that can fill the role of a human, especially in a popular music setting. This will require advances in automated music listening and understanding, new representations for music, techniques for music synchronization, real-time human-computer communication, music generation, sound synthesis, and sound diffusion. Thus, HCMP is an ideal framework to motivate and integrate advanced music research. In addition, HCMP has the potential to benefit millions of practicing musicians, both amateurs and professionals alike."*

As mentioned earlier, my essay is concerned with the OID system. It is not the purpose of this essay to deal with the technical side of the system, however, in order to evaluate its musical characteristics it is important to understand its functionality. The next section provides an overview of how the system developed overtime focusing on musical design.

2. AI Music Generators: OMax-Improtek-Djazz

The following AI generators have been mainly developed in the Institute for Research and Coordination in Acoustics/Music (IRCAM) in Paris, France, with which University of Athens (UOA) has a strong link. In the scope of Improtech 2019 Paris – Athens, the music representation team of IRCAM demonstrated the fundamental principles of these new AI music generators.

2.1. OMax: towards experimental music

OMax is a style modelling system started in 2004. Style Modelling implies building a computational representation of the musical surface that captures important stylistic features in patterns of rhythm, melody and harmony which are then interleaved and recombined in a redundant fashion.[1] It achieves on-the-fly learning and generating from a live source by simultaneously listening to the audio stream of a musician, extracting information from that stream, modelling this information into a formal structure (like a roadmap), then navigating this structure by recombining the musical material to create variations or "clones" of the input. OMax is said to be agnostic in the sense that it has not a priori knowledge or rules to guess or infer any specific analysis of the input based on a particular music theory (Bloch et al, 2008).

Previous research in this field included investigations (in collaboration with Shlomo Dubnov) using dictionary-based models such as LZ compression algorithm IP (Incremental Parsing) and PST (Prediction Suffix Trees). At the beginning it encoded music using MIDI data (Dubnov et al, 2003). Since its first appearance, under the name OMax (Assayag et al, 2006), it included an audio layer (named Ofon) that extracted pitch information from an audio stream (Assayag et al, 2006). MIDI and Audio models build a tree encoding patterns occurrences that allow to calculate probabilities on the continuation of a given pattern (Lévy, 2009). Later, it has been adapted to use spectral descriptors as Mel Frequency Cepstral Coefficients or Linear Prediction Coefficients to build its knowledge (Bloch et al, 2008). OMax's internal visibility structure was

---

[1] As it is explained in the manual of OMax4.5x

developed by Benjamin Levy in 2009 who, by redesigning its architecture, invented new tools to explore such a knowledge and made it modular and more flexible.

Since 2000 -2001, the system is still developing using Factor Oracles for both analysis and generation stages of its improvisation (Nika & Chemillier, 2012). Current research shows that Factor Oracles are the most successful commonly used technique for improvisation (Siphocly et al, 2019).

2.1.1. Factor Oracle Basics

The oracle was initially conceived for optimal string matching, and was extended for computing repeated factors in a word and for data compression (Nika & Chemillier, 2012).

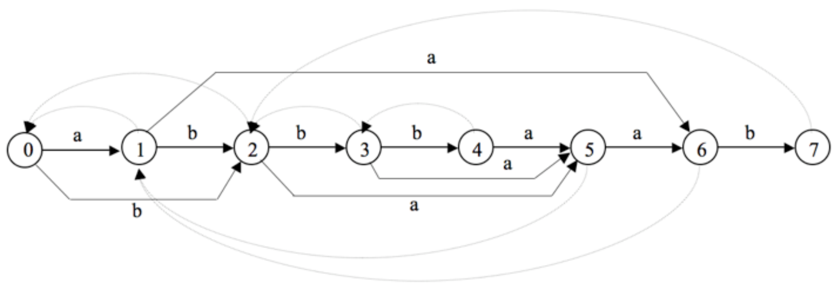

Figure 1: The Factor Oracle for string abbbaab. Suffix links are in dotted lines (Assayag & Bloch, 2007).

Gerard Assayag, pioneer of OMax explains (Assayag & Bloch, 2007):

> *"….The Factor Oracle concept comes from research on string patterns indexation. Such research has applications in massive indexation of sequential content databases, pattern discovery in macromolecular chains, and other domains where data are organized sequentially. Generally stated, the problem is to efficiently turn a string of symbols S into a structure that makes it easy to check if a substring s (called a factor) belongs to S, and to discover repeated factors (patterns) in S (*Allauzen et al, 1999).

Assayag further explains that the above techniques of string pattern indexation share a certain level of description with music, as music, is also sequential and symbolic and the pattern level organization is central to its understanding (Assayag & Bloch, 2007).

The oracle computes a musical stream (corpus) into atomic units referred to as 'states' (Assayag & Bloch, 2007), musical material from the corpus is segmented into different phrases thanks to the detection of silences (Bonnasse- Gahot, 2014), then a learning algorithm builds the statistical model from musical samples and a generation algorithm walks through the model and generates a musical stream by predicting at each step the next musical unit/state from the already generated sequence (Assayag et al, 2016).

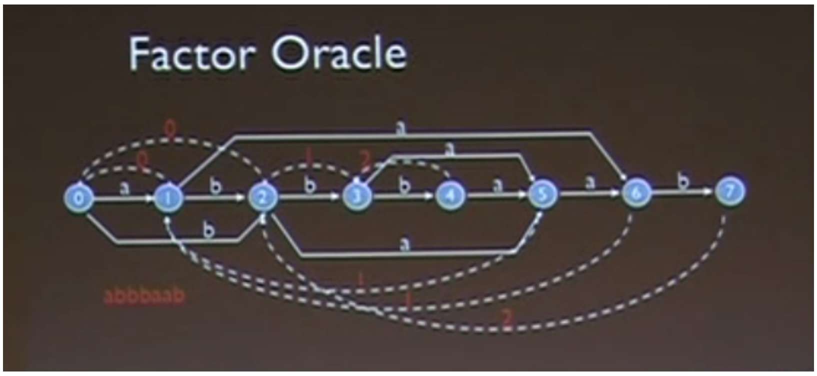

Figure 2: The Factor Oracle. A graph based on symbolic sequences. The dashed backward arrows point and connect the similar patterns.

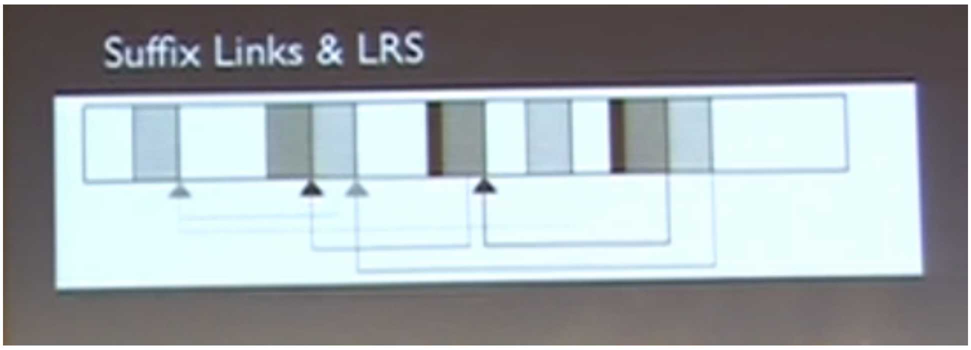

Figure 3: Suffix Links and LRS (length of repeated suffix). A search map of the sequence above and how the patterns can connect to each other, which are similar to which and where they are in the timeline.

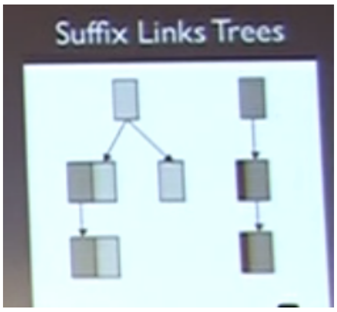

Figure 4: Suffix links trees. All the patterns connected in this these trees have the same suffix, the same ending. These are then used to generate variations. The suffix tree guarantees that the concatenation will be smooth (Levy, 2013).

In OMax5_beta by B. Levy (2013), information extraction of the input stream is done on the basis of three different musical elements: Melody, Timbre and Harmony. Melodic analysis uses the 'yin' object and Statistical Processing. Timbral analysis uses 'MFCC' and 'Vector Clustering' (Euclidean Distance & Weighting). Harmonic analysis uses 'Chromagrams' and 'Vector Clustering' (Cosine Measure).

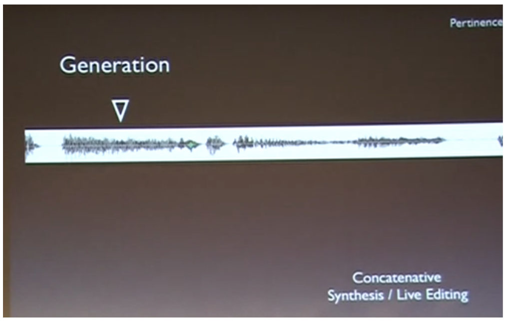

Figure 5: the system reads the segmented data.

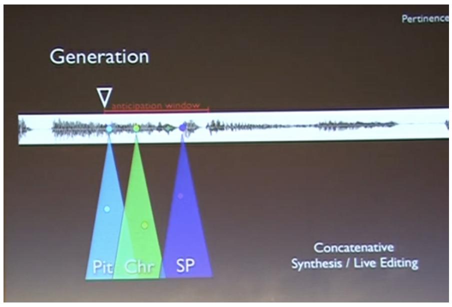

Figure 7: descriptors considered

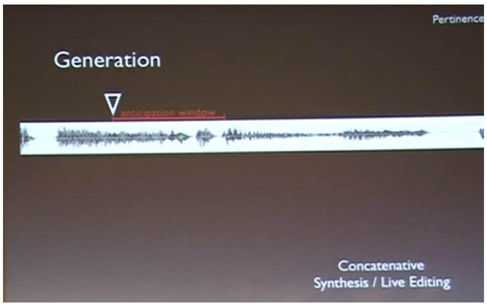

Figure 6: when it is time to make a variation of the learnt material an anticipation window explores all graphs and trees in regards to the descriptors (pitch, timbre, harmony)

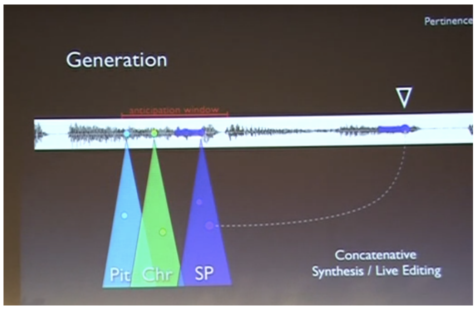

Figure 8: all solutions are gathered and weighted, then the system jumps to a similar pattern and continuous on.

On the above figures 2-8 one can see how a segmented audio stream is indexed and organized according to specific descriptors so as to be used in the reordering of new musical output.

2.1.2. OMax in the context of live performance

According to B. Levy (2013), the major problems of the early version of OMax was the synchronization of the system to either an internal or an external pulse, the ability to follow a time signature, the phrase leadings and the endings. His engineering efforts, in OMax5_beta, introduced synchronization to a pulse as well as synchronization using dynamics and harmony.

OMax was tested with professional musicians in several contexts, specifically duets and bands, with different instruments. This allowed better understanding of the interaction between the clones produced by OMax and the musicians and helped at refining the different analysis stages of the software. Furthermore, OMax was tested in different styles of music and the results showed both the possibilities and the limits. Inferentially, OMax's functionality and musicality is strongly influenced by what is fed, both musically and sonically i.e. length and intensity of musical material and microphone placement in a band situation. Benjamin Levy at al (2012), suggest that:

> *"... Duets of short duration work very well most of the time and some characteristics of purely acoustic improvised duets can be found'....'the computer and its player could really find a place comparable to the other musicians in the group. And the work to build a whole concert program was in almost every aspect similar to the work with an ensemble of only acoustic instruments."*

Furthermore, the musicians in these sessions regarded OMax as a mixed entity made-up by the software and the person controlling it.

Improtek and SoMax projects further developed the listening skills of the system. SoMax by Bonnasse-Gahot (2014) was developed:

> *"...both in terms of melodic understanding, so as to harmonize or provide some accompaniment to a monophonic stream, harmonic listening, so as to make a chorus, and rhythmic abilities, so as to synchronize in real-time with live musicians."*

A great deal of experiments with OMax has been done over the years. A wealth of information, up to 2009, can be found in OMax's official web page. The majority of real-time audio/video examples concern experimental music improvisation freed from constraints of tempo and harmony. A few offline improvisations to test future capabilities, such as Pat Martino's improvisation, had to be inserted by hand in order to follow the tempo and the harmony of the rhythm section. In real-time performance the algorithm navigated the memory and jumped from state to state that shared similar context (Lévy et al, 2012) following pitch and spectral starting points.

2.2. ImproteK: introducing structured improvisation

Second in the chain of OID system is ImproteK. Its improvisations are based on the style modelling performed on live playing or on an offline corpus i.e. a chord progression, refer to as *grid*, making the system able to expand its modelling on harmonization and arrangement. As mentioned earlier, the philosophy of OMax is to turn every single musical event into a *state* in an oracle object. This 'free' real-time navigation thus made it an interactive instrument dedicated to improvisation in a 'free' musical context. Considering that synchronization with a periodic beat is a deep universal of human music perception, Improtek attempted to address the issue of synchronization to a pulse by introducing *labels*, that is musical slices cut per beat based on the current tempo. In that context, it took a metric framework into account by setting the beat as the elementary unit in its acquisitions, restitutions, and generations (Nika & Chemillier, 2012).

ImproteK is a system integrating a rhythmic framework and an underlying harmonic structure in a context of musical improvisation. In Nika and Chemillier (2012) we read that Improtek is similar to OMax as:

> *"…is built on the factor oracle structure taking advantage of the relevant and rich characteristics of this automaton in a musical environment (*Assayag & Dubnov, 2004). *Moreover, it can adapt to a regular beat and produce improvisations following a given chord progression. ImproteK is conceived as an interactive instrument dedicated to performance: its improvisations are based on the style modelling performed on live playing or on an offline* ***corpus****. Combined with pattern reuse techniques, this modelling expands on harmonization and arrangement in a harmonic interaction module."*

A clear differentiation between OMax and ImproteK is given by, Déguernel, K. et al (2017) who say that in ImproteK:

> *"Contrary to Omax, the memory is not based on a sequence of pitches, but on a sequence of musical contents tagged by a label."*

OMax was devoted to free improvisation where chord symbols were not considered; ImproteK was created as a variant of OMax to handle improvisations on chord progressions (Ayad et al, 2018).

2.2.1. Scenario & Memory

ImproteK addresses structured or composed improvisation. The project was initiated by Marc Chemillier in order to combine the style modelling approaches of OMax and the ability to cope with 'beat' and long-term constraints. The software introduces a temporal specification, what is called *scenario*, in the music generation process. Depending on the musical context, this temporal structure can represent several things i.e. a chord progression of a jazz standard, a profile of audio features, or any formalized structure defined on a user-defined alphabet.

As explained in Ayad et al (2018), a scenario is a single sequence and a memory a dictionary of sequences. The letters of memory's sequences are associated with musical fragments, divided in beats, which are concatenated to produce new musical phrases corresponding to the scenario's sequence.

First attempt of ImproteK used MIDI inputs (Nika & Chemillier, 2012), later included audio inputs (Nika et al, 2017). During the learning process these inputs are indexed beat by beat in real-time by the chord labels of the current harmonic grid *scenario*. These fragments, that constitute the *memory* of the oracle, are continuously collected by following the chosen harmonic grid supporting the musical session as a guideline.

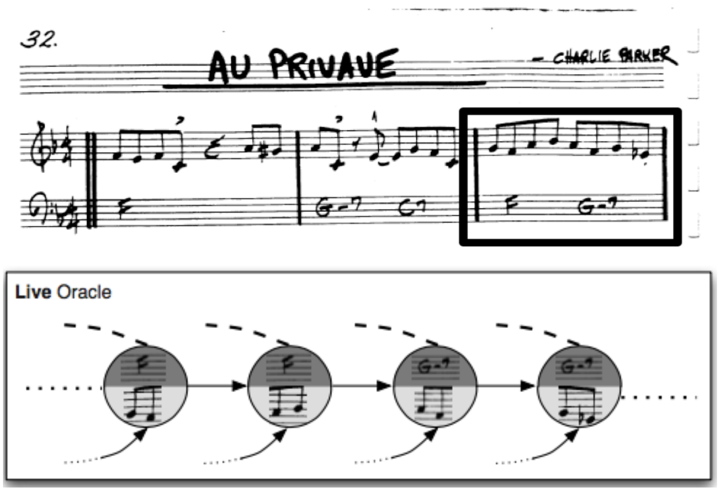

Figure 9: Learning the *live oracle*

For detailed information on the *scenario/memory* model look on (Nika et al, 2016)

2.2.2. Navigation and Synchronization

In the case of a chord progression acting as the given scenario the chord progression becomes a *navigation leader*. This navigation is constrained by the chord labels inserted in the oracle as indexes, a heuristic by G. Assayag. ImproteK marks every beat of the listened live improvisations with the current chord label and then improvises by concatenating the corresponding *beat slices/labels* coming from its musical memory (figure 9). These musical sequences or events are searched and retrieved by an algorithm to be transformed, rearranged, and reordered to create new improvisations. The phrases returned by the improvisation module are labelled sequences.

The interesting with the scenario is that the system takes advantage of this temporal structure. This is done by accessing prior knowledge through anticipation in the generation process. [2] Nika et al (2016) explain:

> *"In this view, reacting amounts to composing a new structure in a specific timeframe ahead of the time of the performance. This point is particularly relevant in the case of "hybrid" improvisation: when the scenario and the memory are different sequences defined with the same alphabet"*

The memory can be made-up by several pieces of music thus reinforcing the hybridization of the system. As we will see later, a hybrid improvisation can be produced by providing the system with a single scenario i.e. a chord progression, and use multiple labelled musical fragments (sequences) from a corpus to generate improvisations.

Another interesting aspect of scenarios described in DYCI2 by Nika et al (2017) is that both scenarios and musical alphabets can be designed by musicians, this way engaging them in a meta-level of composition.

---

[2] that is, to consider the future of the scenario to generate the current time of the improvisation.

To be able to follow the beat, ImproteK integrates a *beat-tracker* which continually estimates the current tempo to be used. This system is described in Bonnasse-Gahot (2010).

Furthermore, a score follower, Antescofo, acts as a sequencer by emitting in real-time the current position in the harmonic grid. For more information on the functionality of Antescofo refer to Cont (2008)

The system also uses a phase vocoder, SuperVP by Depalle and Poirot (1991), whose re-synthesis of sound files can be modified by transformations such as time stretching, pitch-shifting and filtering (Depalle & Poirot, 1991).

2.2.3. Online and offline learning

The aforementioned corpus consists of sourced real-time audio or MIDI data or offline from annotated material like jazz standards (Assayag & Dubnov, 2004). Its learning is performed on three features: the *melodic track* (theme, solo improvisation, etc.), the *accompaniment track*, and the associated *harmonic grid*. The oracle object is built by learning the associations between the three features in the aforementioned slices. The system is able to produce a couple of oracles for every element of the corpus, a harmonization oracle, and an arrangement oracle. My interest in this essay lies on the harmonization oracle which records the sequence of associations between the melodic fragments and the chord labels of the beats it covers. In other words, the harmonization oracle is used to associate a symbolic chord progression with a melodic track. In order to avoid unproductive rigidity which could have been brought by a strict equality criterion equivalences have been introduced. Thus, two states are considered as equivalent if they are indexed by the same notes without taking into consideration their duration, order, or repetition in the beat slice (Nika & Chemillier, 2012).

2.2.4. ImproteK at work

An example[3] regarding the functionality of ImproteK is given by Nica et al (2017):
In this example, the *scenario* provided to the system is the chord chart of the song ‘The Man I Love’, and its musical *memory* is: Billie Holiday singing "The Man I Love",

[3] A link to the project: https://www.youtube.com/watch?v=reJ-SiblCcs

Edith Piaf singing "Mon dieu" and "Milord", Elisabeth Schwarzkopf singing "Mi tradì quell'alma ingrata" (Mozart, Don Giovanni), and "Tu che del gel sei cinta" (Puccini, Turandot).

As we can see memory can be made-up by several pieces of music thus reinforcing the hybridization of the system. For more info on the 'three ladies' project and the hybridization concept one can see (Kapoula et al, 2018).
The result is a virtual mix of the three ladies, singing, and a pianist, Hervé Sellin, improvising, along the chord progression.

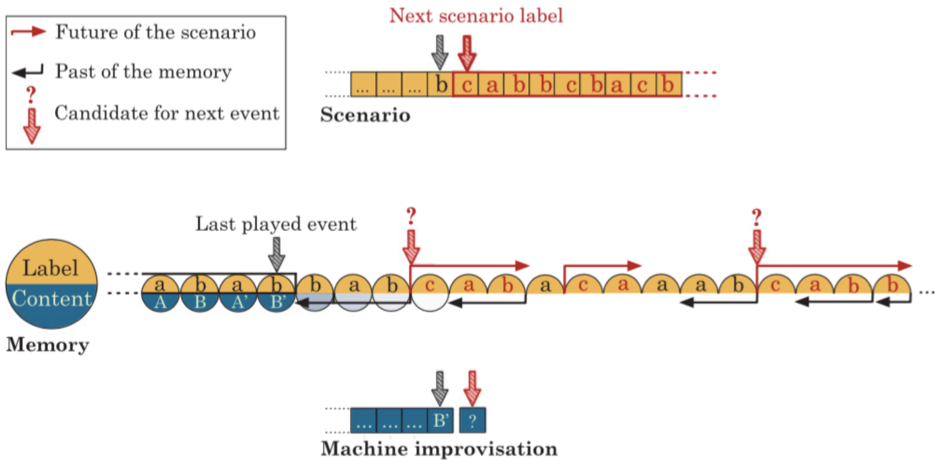

Figure 10: Using the scenario to introduce anticipation in the music generation process.

The general idea of ImproteK is to be part of a band. It is built in a way that can follow both the beat and the harmonic progression. The system needs someone on the computer to be able to pilot it, that is to manage the learning, the generation, and the playing in real-time. The performer has access to material created from the last oracle/phrase learned or from previously learned material. The performer/pilot is in charge of following the chord progression as well as other musical material. Furthermore, ImproteK implements an arrangement module hence it has the ability to act as an accompanist since it can build its oracle with chords rather than melody. Therefore, it can alternately act as both improviser and accompanist

Although not part of the OID series I am examining in this essay, DYCI2 is an interesting extension of the OMax paradigm worth mentioning. Created by Nika 2017, DYCI2 capitalizes on the models OMax, SoMax and ImproteK by *"merging the free, reactive and scenario-based"* music generation paradigms focusing on artificial listening, learning, interaction and generation of musical contents. It aims to take the multidimensional aspect of music into account and also to emulate a collective improvisation situation considering musician's intuition, background knowledge and the influence between them. To further enrich the possibilities of music generation regarding the global form of the tune DYCI2 attempts to exploit the multilevel structure of chord progressions. In order to focus on targeted scientific issues by exploring different idioms and repertoires, DYCI2 project is led in close and continuous interaction with expert musicians (Nika et al, 2017).

2.3. Djazz: towards a consumer product

Djazz is the state of the art of OID system implementing techniques for indexing and creating improvisations using a given chord progression. It relies on a database (dictionary) which is built with musical sequences (audio or MIDI) associated with known chord changes. This database, that constitutes the memory of the system, can be built with either on-the-fly live recordings or with a corpus of jazz solos from great players such as Charlie Parker, John Coltrane and other. This last function constitutes one of the marketing hints of the software. New improvisations are generated by recombing the sequences found within these solos.

Djazz is currently under development. Researchers are focusing on the efficiency of CID (Chord Indexing and Decomposition) algorithm which is presented as a solution to a multiple scenarios queries problem (Ayad et al, 2018), soon to be implemented in OID.

It is worth mentioning here that OID has been studied sociologically and anthropologically by Chemillier, Vigne and Imbert (2015-2018).

	OMax	Improtek	DJazz
Constrained navigation by a scenario	x	v	v
Audio and MIDI	v audio after 2006	v audio after 2017	v
Tap tempo & pulse synchronization	v after 2009	v	v
Video	v after 2008	x	?
On-the-fly recording and file import	v	v	v
Annotated material available to consumers	x	x	v once is marketed

Figure 11: table of functions of OMax, Improtek and Djazz

3. Reflecting on the musicality of OID computer music system: a critical approach

OID is a stylistic reinjection music generator machine guided by a human (Cont et al, 2006) and as such should be assessed. Since it is aimed to improvise as a human (Nika et al, 2016), by analysing musical features of the system my aim is to compare the musical outcomes of OID to that of a human jazz improviser. My conclusions are supported by current research such as the ones found in Déguernel et al (2018), in which professional musicians evaluate the system's musicality and current overviews such as the one by Siphocly et al (2018) that states that music produced by Factor Oracles in not rich enough to satisfy human ears.

OID's developers seem to consider the cognitive foundations of music as the core inspiration of its architecture. This is supported in Cont et al (2006) where they refer to the importance of 'expectation', as laid down by musicologist Leonard Meyer in his book 'Emotion and Meaning in Music'. Thus, they proposed a modelling structure relative to the psychology of musical expectations. Their concept, also borrowing from David Huron (2006), is based on the anticipation mechanism of OID that supports accounts of musical expectation such as: 'dynamic adaptive expectations' and 'conscious expectations', claiming that such modelling constitutes complex musical behaviour such as long-term planning and generation of learned formal shapes (Cont et al, 2006).

A great food for thought is sourced from that text by Leonard Meyer dating back to 1956. According to Meyer, a central thesis of the psychological theory of emotions is: *Emotion or affect is aroused when a tendency to respond is arrested or inhibited* (p 3). A supporting theory by Hebb, found within the same text, suggests that the difference between pleasant and unpleasant emotions is that pleasant emotions are always <u>resolved</u> (p 5), to elaborate, the pleasantness lies not so much at the <u>resolution</u> itself but at the belief of the <u>resolution</u>, it is the control which is believed to exist over a situation.

> *"…The sensation of falling through space unconditioned by any belief or knowledge as to the ultimate outcome, will for instance, arouse highly unpleasant emotions. Yet a similar fall experienced as a parachute jump in an*

amusement park may, because of our belief in the presence of control and in the nature of the resolution, prove most pleasurable…"(p 6)

Motor behaviour, facial expressions, tone of voice and manner of speaking can tell us about the emotional state of an individual. Automatic behaviour is less likely to be emotional. The purpose of emotionally differentiated behaviour is communication. Through a series of non-verbal behavioural signs an individual seeks to communicate his experience. This differentiated behaviour is called 'Emotional Designation'. These signs aim at making other individuals respond in a sympathetic way. And indeed, such sharing does take place in life and in art. This differentiated behaviour is a cultural phenomenon not a natural one. It is based on custom and tradition, it is learned.

In music, experience becomes meaningful when the relationship between the tendency and its necessary resolution is made explicit and apparent (p 7). A tendency to respond, an instinct, is a pattern reaction that operates or tends to operate in an automatic way, in an unconscious level. The order of the pattern reaction is both temporal and structural. This temporality and structure may be upset. That is where it becomes conscious. Such conscious tendencies are referred as expectations. Our conscious mind does not actively seek consequent unless the pattern is disturbed. However, our habits and tendencies 'expect' the consequent relevant and appropriate to itself. Music arouses expectations. The pleasure we derive from style is not an intellectual interest in detecting similarities and differences but an instant aesthetic delight resulting from the arousal, suspension and fulfilment of expectations (p 8).

In particular the variation of periods of tension and release or stability and instability are suggested to provide momentum and define structure in music and thus to be important in to music's sense of emotion.

As you may have noticed above, I have underlined several times the words tension and resolution. Tension and resolution are considered to be fundamental in music creation. I will clearly explain this later in my analysis of OID's performance through the scope of the elements of music.

The text by L. Meyer calls for several interpretations. It can constitute a fertile terrain where supporters and non-supporters of automatically generated music can spend hours debating about aesthetics, ethos and musicality of interactive computer systems such as OID. Recalling Stravinsky's famous quotes (1936, p. 91), alluding constructivist thought, that *"…Music is, by its very nature, powerless to express anything at all…"* (Stravinsky,1936, p. 91) and *"…The trouble with music appreciation in general is that people are taught to have too much respect for music; they should be taught to love it instead…"* (New York Times 1964), the terrain gets interesting especially if the listener is not aware of the ways the music has been produced, was it a machine or a human? The debate gets even more interesting if we further add Douglas Hoftader's (1979) assumption that:

> *"…unpleasant though it may seem, that the "teetering bulb of dread and dream" that nestles safely inside one's own cranium is a purely physical object made up of completely sterile and inanimate components, all of which obey exactly the same laws as those that govern all the rest of the universe, such as pieces of text, or CD-ROM's, or computers."*

This debate regarding the interactive improvising machines is a long standing one. George Lewis (2018), one of the pioneers and warm supporter of human-computer interactivity in his efforts to answer the question *why do we want our computers to improvise?* invokes notions of complex matrix between humans and machines, the possible and the impossible, new forms of freedom, new modes of everyday life as well as philosophical, social, political, cultural, technological and music-theoretical interests and concerns. On the other hand, the non-supporters are resisting to the *spontaneity mapping* efforts of George Lewis and the likes insisting that the mindfulness of music improvisation should remain a mystery and interests such as the ones above should not be encompassed.

Contrary to George Lewis who suggests that current systems, in its current incarnation, perform well, Gary Marcus (2019-20), an AI and cognition expert and CEO of robotics company Robust.AI says that AI isn't very smart "*we're just scratching the surface of intelligence."*

His assertion is that Deep Learning without Deep Understanding won't get us anywhere near human-level intelligence. Although, Gary Marcus refers to GPT-2, a text generator, the same may apply to music generators too. After all, Marcus ideas reflect

Douglas Hofstader about understanding oneself. "*How do you make a search engine that understands if you don't know how you understand*? *AI has become too much like the man who tries to get to the moon by climbing a tree"*. Deep Blue, the IBM supercomputer that won over Gary Kasparov won by brute force. "*So What*" says Hofstader, "*does that tell you how we play chess*?" (Somers, 2013). AI may need to shift from deep learning to deep understanding so as to become conscious, self-referent. At least in theory, this is possible through the concept of the *Strange Loops* by Hofstader. Till today, although exciting and promising in music generation, AI is a fairly limited tool compared by many as having a toddler's brain.

3.1. Performance outcomes of OID and the Elements of Music

> *"As should be clear to any musical reader, assessing a music generator in an objective manner, if not impossible, would set along disputable measures of goodness. On the other hand, in most music practices and styles, what is considered as wrong can be constituted as a feature depending on the context."*

The above statement was made by OID developers in (Cont et al, 2006). Provided that we are discussing a system which functions idiomatically, specifically in the jazz idiom, I believe there are several problems with this statement. My point will be made clear below.

Declaration: the text on the following sections constitute assumptions based on: my own experiments using several demo versions of OMax and DYCI2, research through available resources such as published scientific papers, videos and audio recordings published in the corresponding web pages. The system is still under development (Ayad et al, 2018) and is employing professional musicians to evaluate its musicality (Déguernel et al, 2018).

Following, I present my musical analysis of OID in regards to the elements of music.

3.1.1. Beat, Meter, Rhythm, Groove, Silence

OID's synchronization has been addressed over the years through ways such as the one described in ImproteK. The system seems to be able to both follow an internal and an external pulse either through a beat tracker or through *bang* tap-tempo gestural sensors

(Bonnasse-Gahot, 2010). Moreover, beat phase of the segmented musical events is annotated within the beat structure and coupled with micro timing mechanism preserving the *feeling* and the *swing* (Bonnasse-Gahot, 2014). In the case of automatic beat tracking however, serious mistakes arise i.e. losing track due to slightly offset events, constituting essentially the groove of the tune, or miscalculating half and double times. This is a problem in general in MIR research and is evident in i.e. M4L beatseeker plugin and in Spotify's tempo calculator. Humans and beat trackers are using different processes to identify the beat and these are not always consistent. Tapping, especially foot tapping, is favoured over automatic beat tracking as most reliable. However, distraction during performance may prove problematic in defining precisely the beat (Dannenberg, 2012).

To my knowledge, throughout the development of the project, there is no sufficient research regarding OID's response to beat accentuation and time signature. These features are style defining in jazz performance and composition. Backbeat accentuation and syncopation is at the heart of jazz. For example, in simple meters i.e. 4/4 time, norm for most of jazz repertoire, beats 2 and 4 are important not only in regards to velocity but in chord and scale tones placement of melodic material. This is true for both traditional and modern styles of jazz. An improviser, for example, may place chord tones on beats 1 & 3 and non-chord tones on beats 2 & 4 (figure 12). The opposite may also happen. Having control over this placement can reaffirm a tune's form or provide a rhythmic counterpoint to the form plus it extends a musician's pallet with colours and suspense (Berliner, 2009, p. 318).

Figure 12: chord tone placement on beats 1 and 3, Bert Ligon on 'Connecting Chords with Linear Harmony'

As Larry Gray, one of the finest jazz bass players explains in Berliner's book 'Thinking in Jazz' (2009, p. 317).

> *"You are always phrasing into downbeats. In 4/4 time for example players tend not to think in terms of the beat grouping "one two three four" but in terms of*

the beat grouping "two three four one". In this sense their lines are always moving ahead, jazz is over the barline music."

Moreover, when compound or odd meters come into play, also found in jazz repertoire, then several problems arise. First the division of the meter is laid out in a specific way i.e. 2+2+3 for a 7/8 odd meter (figure 13). In such case, the accents of several elements of music i.e. melody, rhythm, harmony, timbre, dynamics, fall onto specific places within the meter. Second, and this is true for simple meters too, multiple time signatures often coexist within a meter (cross rhythms, rhythmic hemiolas) i.e. 3 over 2 or 5 over 4 etc which is achieved by accenting the beginning of a series of notes i.e. a developing motif based on the notes of a scale or of triad-pairs in groups of 5. It will be interesting to see how, if at all, OID and similar systems will implement these techniques. Will the machine be able to jump from one state/level to another with respect to the aforementioned characteristics of jazz? In this direction the work of Srinivasamurthy et al (2014) regarding rhythm in such diverse styles of music such as Turkish and Indian may prove helpful.

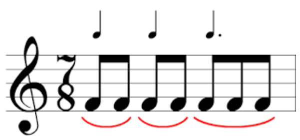

Figure 13: An example of odd meter with accents on beats 1, 3, 5.

Rhythm, a defining characteristic of jazz, is improvised during performance. An example may be a guitarist improvising and a drummer following the guitarist's accentuations (kicks) or playing cross rhythms creating an intense multi-sensual experience for both the performers and the audience. OID plays the segments with no real-time listening capabilities. This lack of real-time *sense* also affects the groove which, even if it stays constant throughout, slight fluctuations of live performance tempo will cause misalignment of previously recorded material, especially if the operator is not fluid with keeping up with the tap-tempo function. In the video[4] titled 'Katsepy-Save the Earth' at 4:27 the lack of groove is more than evident. However, on this video[5] 'Katsepy- Ho any an-tanana' at 3:27 the groove sounds tighter, for a short

[4] https://www.youtube.com/watch?v=tsTI2M0OBWg&t=217s
[5] https://www.youtube.com/watch?v=NTWjYJj0LX8&t=147s

period at least. How easy is for OID to keep up with different degrees of "swing feel" i.e. a "laid back" and "ahead of the beat" stylistic approaches during performance?

Silence, for many musicians and educators, is a stand-alone element of music. Indeed, without silence how one could appreciate sound organized in time? When playing with OID it is up to the human operator to apply pauses during performance, my experiments showed that the system doesn't pause by itself at any time.

3.1.2. Harmony

Not a surprise, MIR's research on Automatic Chord Extraction (ACE) performances has reached a glass ceiling (McVicar et al, 2014). Machine estimation of complex chords as seen in Jazz, is a daunting task. This is also true for a human being learning to recognize jazz chords, let alone progressions of chords. To date, with scores of up to 73% of accuracy in popular music songs, ACE systems are only able to feed application areas such as mood detection, song analysis, music recommendation and structure analysis, not used into real-time music making instrument. Hence the use of scenarios by OID.

A scenario in OID, as explained earlier, can be a chord progression which becomes a navigation leader of the oracle. This navigation is constrained by the chord labels inserted in the oracle as indexes. From a jazz performer's point of view, feeding the system with a static chord progression presents some problems. First, regarding credibility, where was this chord progression sourced from, was it from a Real Book, from an ACE transcription or from a human transcription, of which recording? Second, chord progressions in jazz constitute just a lead sheet meaning that this can be altered during the performance i.e. a D7 may become D7b9 or completely replaced by a substitution chord such as Ab7alt (its tritone substitution) etc, put simple, chord accompaniment in jazz performance is also improvised. Third, once the harmony is altered by the accompanist, the soloist may, or may not, follow this alteration in his/her improvised line. Also, through her line, may suggest a different chord that the accompanist may or may not follow. This kind of musical interaction between musicians, live on stage, is common practice and a defining characteristic of jazz performance.

Déguernel et al. (2017), borrowing from Noam Chomsky's concept of phrase structure grammars[6] (Chomsky, 2013), considers possible variations of a tune's structure by using generative grammars to create a hierarchical analysis of a chord progression, thus creating a multilevel progression, that is, hierarchically structured multiple instances of a given progression (figure 14). The equivalency of these chord progressions is analysed and validated by a professional jazz musician and the actual contents of the multi-level progression are automatically learnt on a corpus.

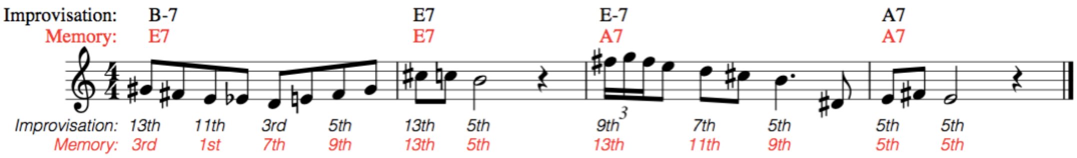

Figure 14: First four bars from the bridge of an improvisation on *An Oscar for Treadwell* (rhythm changes form). The 'Improvisation' lines shows the chord and degrees played in the generated improvisation; the 'Memory' lines shows the chord and degrees on which the different parts of the melody used to generate the improvisation were actually learnt (Deguernel et al, 2017).

My understanding, although in their paper they refer to this as "*changing voicings*", is that through this method the system is able to improvise on a progression of chords, including substitution chords, as long as they share a similar role in the scenario as previously known chords. This is achieved because the multi-level labels from the memory and the multi-level progressions from the scenario are equivalent. This achievement is addressing the second aforementioned point I make: the system through adaptation is able to improvise on previously unmet chords. It would be very interesting to see if a generated improvisation considers a system's capability to recognise chords on the fly.

Recent research focuses on the introduction of musical knowledge in the representation, the improvement of the models towards more complex chord alphabets and the development of more adapted evaluation methods. Convolutional Neural Network (CNN) architecture comes into play aiming at developing ACE (Carsault et al, 2019).

---

6 *Phrase structure grammars* are a type of grammar presented by Noam Chomsky based on constituent analysis, that is to say on a breakdown of linguistic functions within a hierarchical structure.

This direction sounds encouraging as it will develop the listening capabilities of systems such as OID.

Although not the theme of this essay, is worth mentioning here that in 2012 ImproteK implemented an arrangement oracle that, also, used equivalency in order to associate an accompaniment track to a symbolic chord progression (scenario), thus able to improvise the accompaniment (Nika & Chemilier, 2012).

3.1.3. Melody

Traditionally, jazz musicians improvise within a harmonic framework, this means that they play a series of notes which *fit* the given chord. Within these series of notes several techniques may take place. The ones listed below are some popular melodic techniques whose musicality depends on beat placement, plus they strongly address the tension-resolution concept, a concept associated with universal music-making:

1. Voice leading: 7 to 3 or/and the opposite, meaning that the melody moves, over the bar-line, by step from the 7th of the current chord to the 3rd of the following (figure 15).
2. Enclosures: enclosing either a scale or a chord tone. It can happen anywhere within the beat but is most profound on beats one and three. This often involves *forward motion* where the melody starts on an off-beat usually with a chromatic note (figure 16).
3. Bebop scales: As a rule of thumb the added chromatic note must fall on an offbeat (figure 17).
4. Resolution colors: resolving by step or a skip of a fifth suggests strong resolution. Other motions provide weaker resolutions signaling the arrival of the next chord less definitely (figure 18).
5. Contour: emphasizing horizontal melodic constructions. It can carry listeners over the bar-lines towards longer-range goals (figure 19).
6. Endings: phrase endings i.e. resolving on the 6^{th} of the chord on the off-beat of two (figure 20).

Figure 15: 7th to the 3rd, Charlie Parker on 'Donna Lee'

Figure 16: chromatic enclosure, Michael Brecker on 'Softly as In a Morning Sunrise'

Figure 17: use of bebop scale, Charlie Parker on 'Donna Lee'

Figure 18: resolving by step, John Coltrane on 'Lazy Bird'

Figure 19: ascending contour, Michael Brecker on 'Softly as In a Morning Sunrise'

Figure 20: resolving on the 6th on the off-beat of two, Charlie Parker on 'Donna Lee'

The above are just a few examples of jazz improvisation techniques governed by conventions within the idiom. They are part of what is called 'the language of Jazz', both traditional and modern. Melodic devices that efficiently and stylistically address the tension and resolution concept. Keeping, or not keeping, with these conventions should be under the control of the musician. Melodic patterns over chord changes are defined by both rhythm (placement within the meter) and direction (ascending, descending, arch, inverted arch, stationery). These conventions constitute fundamental melodic construction developed over hundreds of years (Crook, 1991). To my understanding, the aforementioned techniques are not addressed by the OID system. OID finds good transition points according to available descriptors (Malt & Jourdan, 2008) chosen by the human handling the system or a scenario (Nika, 2017). Reinjecting the captured material in random places within the meter, process defined by either the scenario or the descriptors, does not produce a jazz line governed by successful tension and resolution. Watching the video[7] titled 'Patchwork of Autumn Leaves: Cellin, Mahler, Puccini, Mozart' (2:30) provided by Nika in his YouTube channel experiment with ImproteK one can hear such 'unmusical' jumps happening real time. The contour as well as the tension and resolution of the melody is, more than often, compromised. My analysis is further supported by professional musicians collaborating on the development of OID (Déguernel et al, 2018). Double bassist Louis Bourhis evaluating generated improvisations on two tunes: 'Anthropology' and 'Donna Lee' said that the improvisations were *patchworks* of Parker:

> *"…Harmony makes sense in a continuity. . . . At the moment, it doesn't take that into account, or it is juxtaposing them in a random manner. We don't really hear harmony. We hear note after note, or phrases after phrases. And even inside phrases, there is not necessarily any harmonic sense…"*

In justifying their thesis Nika et al (2016) make a parallelism of OID with several forms of music making such as jazz, rock, blues, basso continuo and Indian raga. Motivated by the above statement and considering that jazz is also a transcultural idiom, I comment that in some styles of music including Indian Raga, Greek folk and Turkish maqam among others, the intonation of pitch is changing according to the contour of the melody i.e. a flat is 'flatter' than the flat of Equal Temperament System when the melody descends (microtonal music). Such style-defining characteristics are yet to be addressed in the development of machine improvisation systems.

---

7 https://www.youtube.com/watch?v=dNcK3bRnbv8

3.1.4. Timbre

Compare with other algorithmic improvisers OID is one of a few systems that handles audio. Through this research it was made clear to me that this is an achievement in itself. However, in both injection modes *live* and *play-file*, fluctuations of tempo raise the need for time-stretching. In *live* mode this can happen due to possible fluctuations from the moment of the recording to the moment of the performance and in *play-file* i.e. an annotated solo by Charlie Christian, tempo will have to be adjusted to accommodate the current performance's tempo. Furthermore, pitch shifting, which is used to accommodate new tonal centres, will be needed for *play-file* mode to transpose the file to the current key and for live injection mode when, for example, Dual_SP & SVP players are used to harmonize the current improvisation as seen in OMax4x[8].

Time stretching and pitch shifting applied to an audio signal introduce unavoidable small latency as well as artifacts in the transformed sound, compromising the produced audio resulting in quality loss therefore affecting timbre (Liuni & Röbel, 2013). Nevertheless, it seems that algorithms like the one produced by IRCAM analysis & synthesis team (see Lab TS 1.0.11) perform well even with extreme time-stretching (Průša & Holighaus, 2017), (IRCAM Analysis Synthesis Team p.1680, 2020). Moreover, as with any audio fed into an audio editing software i.e. Max and Ableton Live, the timbre can be creatively manipulated with the addition of FX.

3.1.5. Dynamics

One of the finest qualities of music composition and improvisation is dynamics. They strongly relate to the immersion of feelings. OID is not made to produce new music but rather to reproduce what has already been played as stylistic reinjection. Therefore, the handling of dynamics is based upon the musician feeding the system and the human guiding it. As such, the dynamics of the produced improvisation of OID present several issues. Careful feeding of OID has to be considered in order to avoid extreme dynamic changes due to uncontrollable and unpredictable jumps in the memory. Current research (Déguernel, 2018) seems encouraging but is not clear whether this multidimensional improvisation will have an effect on dynamics. If so, the improviser feeding the system will eventually gain more freedom as to what to play when feeding the system.

[8] explained in the manual of OMax4x

3.1.6. Texture

Texture cannot be analysed in depth yet as at least OMax is not able to produce other than strict intervallic homophonic counterpoint through the SVP players. ImproteK and Djazz are not available for evaluation.

3.1.7. Structure – Form

On this section I am using the concept of story-telling in jazz improvisation, a term with long history of prominence within the style, to draw useful conclusions in regards to the structural performance of improvisation of OID. Storytelling is one of the finest and probably one of the latest endeavours of jazz musicians, it strongly demonstrates individuality. The concept in itself is self-contradictory: how could stories be told without words? plus, it may interpret in several ways, is open-ended (Barrett, 2000, p. 19). To get a better understanding of the term, below, I display several definitions given by the masters of jazz improvisation in the context of two descriptors: unfolding oneself and mood & emotion.

Unfolding oneself. This is strongly related to continuity, development and cohesion within a *solo*. Great musicians demonstrate these in their improvisations. They introduce musical segments, motives and phrases which are put together in such a way so as to make a story. Throughout their improvisation they juxtapose ideas introduced earlier, like a 'Leitmotif' that have borne in mind all along (Paul Wertico in Bjerstedt, 2014).

> *"From the first note that you hear, you are responding to what you've just played: It's like language: you're talking, you're speaking, you're responding to yourself. When I play, it's like having a conversation with myself."* — Max Roach in Barrett (2000).
>
> *"Parker played Charlie Parker, not bebop. People say he played bebop, but he played himself."* – Marsalis and Hinds in Bjerstedt (2014)

This technique is not new. Early jazz players such as Louis Armstrong and Sidney Bechet and swing players like Lester Young structured their improvisations within multiple choruses. On the first chorus they would interpret the melody and on the next they would take expressive liberties by varying it and then return to the melody or play single not riffing patterns. Shaping is essential to both early and contemporary players and as Buster Williams says:

> *"Calm down and start from the beginning. "It's also like playing a game of chess. There's the beginning game, the middle game, and then there's the end game. Miles is a champion at doing that. So is Trane. To accomplish this, the use of space is very important—sparseness and simplicity—maybe playing just short, meaningful phrases at first and building up the solo from there."* – Buster Williams in Berliner (2009)

Great players demonstrate restraint. They leave space by starting simply and lyrically and if intensity builds their ideas become more complicated and longer. This way they create a destination point. Kenny Barron explains:

> *"The way I look at it is that you're going to start down so that you have somewhere to go. It can build to different points in different parts of the solo. It's hills and valleys. That's what it is anywhere. It could happen in different spots within the tune at different times."* – Kenny Barron in Berliner (2009).

Going back to the point I made earlier about tension and resolution, Sonny Rollins suggests that a beautiful story is made with harmonic resolutions:

> *"It's emotional. ...somebody wrote that what I was doing in certain song was asking a question and then answering the question. I think he was talking about harmonic resolutions. So that would be sort of what I think telling a story might be: resolving a thought."* – Sonny Rollins in Bjerstedt (2014, p. 23).

<u>Mood and Emotion.</u> The mood and the perceived emotion of a tune affects the chosen rhythmic intensity and tonal material of the soloist. Moreover, chord progressions affect the dynamics and the density of melodic material.

> *"A piece's emotional associations commonly influence an artist's rhythmic approach or selection of tonal materials, in the latter instance suggesting,*

> *perhaps, an emphasis upon blues-inflected melodies rather than brighter, uninflected melodies or upon tense rather than relaxed harmonies….”* - Emily Remler in Berliner (2009)

There is no ‘one size fits all’. Every human expresses themselves in a different way. Extreme fluctuations of feelings are evident within the same tune played by different performers (Roberta Baum in Berliner, 2009). This may be due to the balance between emotion and intellect both of which are partners in the conception and expansion of ideas (Harold Ousley). Though some players strictly invoke emotion: *“talking straight out of your heart”* – Jan Lundgren in Bjerstedt (2014)

The idea of story-telling is approached here by the masters. Telling stories when improvising jazz entails large, important and difficult questions (Bjerstedt, 2014, p. 15). The notion that “if jazz tells no story it is simply not good” is widespread among jazz masters (p. 17, 18)

> *“the problem today is that good improvisers are so rare. There are many people who can make sense out of their improvisation, but very few are really saying anything.”* – Stan Kenton (Bjerstedt, 2014, p. 18).

Put aside the emotional aspect of the term, as we already mentioned we are assessing a machine, evaluating OID’s capabilities of structured improvisation becomes difficult. The concept calls for metaphors that are affected by emotional, social, national, racial and elitist characteristics. However, throughout my research regarding OID I haven’t come across a point that addresses storytelling, in other words, the long-term structure of improvisation. The issue of structure in AI music is currently one of the major problems not only in OID but in other major projects such as Magenta by Google (Engel & Roberts, 2018) and even on the text generator GPT-2 by OpenAI (Marcus, 2020). Having said that, from my experience as both a musician and an educator structure is one of the most challenging elements to grasp, hear and perform/compose in a music piece.

The self-referential point Hofstader makes through Bach’s music may prove beneficial here. OID indeed functions as a self-referential machine as it reinjects the learned material. *Strange loops* may be applied in the construction of a model where tangled hierarchic layers of motifs will reference one another leading back to the initial motif in a *Strange Loop*. Sonny Rollins is a master of this technique.

Vassilakis' et al. (2019) concept, once it finds itself into a working module, may also contribute reaching an effective story telling by machines through a syntax that connects musical segments, licks and phrases.

4. Discussion

Through this research I got the belief that, to couple Gary Marcus, indeed machines are taking over the world but they are still dump. Improvisation capabilities of AI music generators have increased but in expressive musical terms are still far from being able to reproduce a musical and intense, human-like, improvisation performance. Considering the ideas by Douglas Hofstader as laid out in his book Godel-Escher-Bach, it seems that artificial intelligence systems could benefit by investing more into deep understanding as opposed to deep learning. To my understanding, OID, by using factor oracles, can walk towards this direction as is one of the few, if not the only one, AI system that both doesn’t need a large amount of data to create its improvisations plus it considers Chomsky’s, a nativist philosopher, point of view about structure. This is in contrast with empiricists such as GPT-2 co-founder Ilya Sutkever from OpenAI's who support big data management.

Stylistically, OID stays loyal to the characteristics of the idiom as it reinjects the learned material. This make it sound jazz-like but further developments are needed to produce a jazz line governed by conventions. Relying on chance doesn’t guarantee a musical let alone an emotional experience. On its own right, although is not a *jazzer*, OID is an artistically credible machine as it is proposing something new into the world of idiomatic performance. Through continuous research this approach of melodic construction may develop in different directions. However, the tension and resolution concept should remain vital.

Constant ongoing transcription of solos produced by OID aids the musical analysis and comparison to that of a human improviser. Listening and evaluation sessions with professional musicians are vital to its further musical development. An experiment with music produced by machines vs humans will be interesting to see how people react i.e. do they realize that this is a machine, how it makes them feel? etc. Moreover, a questionnaire can be produced for relevant participants such as musicians already involved with the project, electronic musicians handling the system and researchers developing it.

Continuous research in the areas of expressiveness, creativity, ethos and aesthetics from a performer's point of view will help modelling more effective machines.

It's also crucial to look on the system from a performance studies view and examine the *Spheres of Performance* as laid out by Richard Schechner (2017). For example, what does AI performance do, what are the goals in terms of:

1. Healing people.
2. Dealing with the divine and the demonic, dealing with worlds that are invisible to people.
3. Teaching or persuading.
4. Making or changing identity
5. Fostering community. Creating solidarity among groups
6. Creating beauty
7. Entertaining

Furthermore, what is the performance text and behaviour text of OID and other AI generators? How is the embodied or non-embodied behaviour expressed? Performance Studies is an interesting area to expand upon when dealing with music generators. It may help define clear lines in regards to the ethos of AI.

It will also be interesting to examine the generated music from a cognitive point of view. Rhetorically, is this kind of music making a left brained process? How it can engage the right brain and stimulate emotion for either the creator or the listener? On that note, one can claim that OID's generated melodies fit within the jazz idiom. However, I would argue that OID behaves the same way as a toddler jazz improviser who is not really saying anything through his improvisations. Few *jazzers* are able to touch one's soul and create music that reflects something.

Interacting with a machine in music making is controversial among the music community. Approaching this relationship, one need a clear mind to realize what is happening. Music generators should be seen as such and not compared to a human musician. Only then a new form of interaction may produce new music. Jazz musicians building music on top of each other, collectively improvising in synchrony, in a primal level of communication is not yet possible with music generators.

5. Conclusion

I have provided a general functionality description coupled with a musicological analysis of the system OID in the context of Jazz. Such analysis and evaluation of machine generated improvisations based on the elements of music are proving valuable and may be applied to similar systems and expand to other scientific areas.

It is explicit that the major drawback I pointed in OID is that of tension and resolution concept. A development point for a more musical melodic construction may be to use MIR techniques to find the chord tones, the non-chord tones and chromatic notes within a musical phrase annotated by a chord symbol and push the system to place these at the deliberately chosen beats within the meter following the given scenario (chord progression). An algorithm such as the one proposed by Konstantelakis (2019) may be used towards this goal. This can be further enhanced with an adjustable parameter i.e. from strict chord tone application to non-chord and chromatic tones placed on the chosen beats. Another idea may be to use transposition i.e. SuperVP, of the melodic material to smooth the transition from one bar to the next this way avoiding random 'jumps' when a new chord arrives. This may be achieved with a restriction of 'jumps'. Both, transposition and restriction will work towards effective contour. Furthermore, techniques of improvisation as methodically and meticulously laid out in Crook (1991) can be used as a guide to design modules that drive the system towards the production of effective lines.

Musicians play with OID should be aware of the drawbacks in relation to the groove/ beat. One needs to know what to feed the system so it can take advantage of its capabilities. There are limitations. The human operator ideally should be a musician as the musicianship of the human handling the system is inevitably affecting its performance.

A human can play side by side with a human and a machine can play side by side with a machine. However, to date, a machine still needs a human to play side by side with a human. This human-machine coexistence constitutes a general admission of MIR.

6. Performance Analysis

My performance is made up of two tunes. The first is “Chameleon” by Herbie Hancock from the album *Head Hunters* and the second is a duet for Omax & Saxophone. “Chameleon” is performed solo with live looping and electronics and the second as a duet in which I am piloting OMax.

6.1. Discussion

Production and performance practices are trending towards each other. The lines become more blurred. New models for performance and a significant flux now exists between the two spheres of musical activity which is seeing remarkable new practices emerge.

In this context, I undertake a live, improvised construction of a piece as a performative act. Recording process is thus revealed on stage in real time and I perform a process that would have once been confined to the recording studio.

The incorporation of recording studio practices in live performance is most widely known in the tradition of Jamaican ‘dub’ by practitioners such as King Tubby and Lee ‘Scratch’ Perry dating back to the late 1960s. Outside the popular music tradition, live ‘performance mixing’ can be found much earlier within the electro-acoustic music tradition from the late 1940s onwards in the work of prominent composers such as Bernard Parmegiani (Knowles and Hewitt, 2012).

I approach the act of mixing in much the same way as an instrumentalist might approach performance on a conventional instrument. For this task, I am using a dedicated MIDI controller, working its faders, pots and switches in an overtly instrumentalist fashion to improvise a mix of the source multitrack recording.

I also use MIDI, for real time control, storage and recall of a range of sonic parameters. ‘In the box’, I use software plug-ins optimized for use in live performance. I also synthesized and processed my own patches and a range of parameters that I could use live via gestural control, automated against time, or triggered as a sequence by specific performance events. Software synthesizers and samplers are running from within my laptop in a DAW.

For *Chameleon,* as a performer I will not only record on stage but I will integrate this with the process of composition and arrangement. This provides an opportunity for the audience to not only experience me recording a take, but also to catch a glimpse of my studio process more broadly. They will see me at work, building and arranging a track, in much the same way as I might work in the studio, but in this case organised and framed as a performance. Furthermore, on the experimental, improvised music of Omax & Saxophone duet, although a graphic score will provide an elementary structure and I will be controlling many of Omax's parameters, I will be dealing with an element of surprise, resulting from both the machine and the human playing the saxophone, which has to be addressed and manipulated live.

6.2. Chameleon

6.2.1. Musical notes

The key centre of the tune is Bb Dorian. The full bassline consists of broken arpeggios anticipated by walk-ups made of chromatic notes. The rhythm of the bass-line pattern is similar to the 3-2 *son clave* pattern. A characteristic difference is the three eighth-notes anacruses, which creates an interesting, ubiquitous tension. The way that the bass-line is being played on a synthesizer makes the perception of the downbeat ambiguous. This tension lasts until the human beat box starts, which resolves the riddle.

The snare rhythm has a hit every three sixteenth notes. This figure is known in Afro-Latin music as *tresillo*, popular within all styles of American popular and vernacular music. *Tresillo* also forms the front half of the equally ubiquitous 3-2 *son clave.*

Figure 21: bass and drums on the first section of "Chameleon"

The tune is built entirely on a two-chord vamp Bbm7 (Im7) and Eb7 (IV7).
The various keyboard, guitar and horn parts use the B-flat minor pentatonic scale to create polyrhythmic riffs, layers that build gradually one upon another.

6.2.2. Background

Released on October 13, 1973 on Columbia Records by the American pianist and composer Herbie Hancock *Head Hunters* became not only Hancock's most successful album, but one of the bestselling jazz albums of all time. Apart from his strong classical and jazz background Hancock cites James Brown and Sly Stone as major influences in the creation of *Head Hunters*. The album's centrepiece is "Chameleon'" a 15-minute funky meditation which demonstrates Hancock's increasing fascination with electronic keyboards such as the ARP Odyssey synthesizer, the instrument responsible for the song's distinctive bass line.

6.2.3. Technical notes

On the technical side, I am using a DAW employing built-in and third-party plug-ins. There are nine tracks in total of which one is an instrument track, six audio tracks and three buses. The control of the performance is achieved through three MIDI controllers, a foot controller, a dedicated mixer & launch controller and a keyboard controller that mainly controls the Virtual Instrument. The guitar plugs on a D.I. box and a wah-wah pedal before the soundcard, from there is treated with effects within the DAW.
The session is programmed in a way that demonstrates structure with parts coming in and out of the arrangement at the push of a button. This approach goes beyond simple, time consuming looping techniques. Part of the technicalities is pre-programmed into the DAW so as to avoid delays on layering and to aid the manipulation of structure. All parts are recorded, mixed and manipulated real time.

6.2.4. Performance structure

This Jazz Standard is one of a few modal tunes that I can reproduce using looping techniques. My arrangement imitates the original tune in terms of structure. The different layers are gradually built one upon another using synthesized sounds, guitar and beat box. Once the rhythm section is complete a solo guitar takes place. At the end of the solo a soundscape is created using the already recorded sounds. The texture of this last section alludes Coltrane's "sheets of sound" in that it tries to mix several elements of the tune within the same vamp using live electronics which include multiband filtering, freezing, looping, scrolling and stretching effects. Rhythmic entropy is gradually introduced through granular synthesis creating a complex mass of sound.

6.3. OMax & Saxophone Duet

The key feature of this duet is the interaction between human and his/her musical "clone" produced by the machine. OMax captures part of the melodic material of the saxophone player then extracts information from that stream, modelling this information into a formal structure (like a roadmap), then navigating this structure by recombining the musical material to create variations or "clones" of the input. I will use this performance to draw further conclusions in regards to the aesthetics of human-machine interaction, musicianship and technical issues in live performance.

6.3.1. Technical notes

I am using the OMax system developed at IRCAM, France. The system runs on Max platform. The different modules are MIDI mapped within Max and controlled with an external MIDI controller. Both OMax's and saxophone's audio are routed to the DAW within which several effects are applied on both during the performance.

6.3.2. Performance structure

The structure is governed by a graphic score by Carl Bergstrom-Neilsen, a Danish composer and improviser. The original score is a graphic representation of a journey from heavy dark sounds to light sounds. Players progress at their own rate through the spiral representing this change, loosely coordinating with the other players. Our performance takes the opposite direction moving from the centre towards the exit or from lighter to darker sounds. The performance starts with the saxophonist freely improvising using a mixture of mainly slow-moving melodies with few fast ones. OMax captures part of the improvisation in order to build its first oracle. Several areas of the oracle are then chosen to be reinjected. This is the first stage of the interaction. The tune continues to develop through the same process changing the injected and the reinjected material. An experimental process that can be compared to a Playback Theatre where actors acting out a novel story without rehearsal.

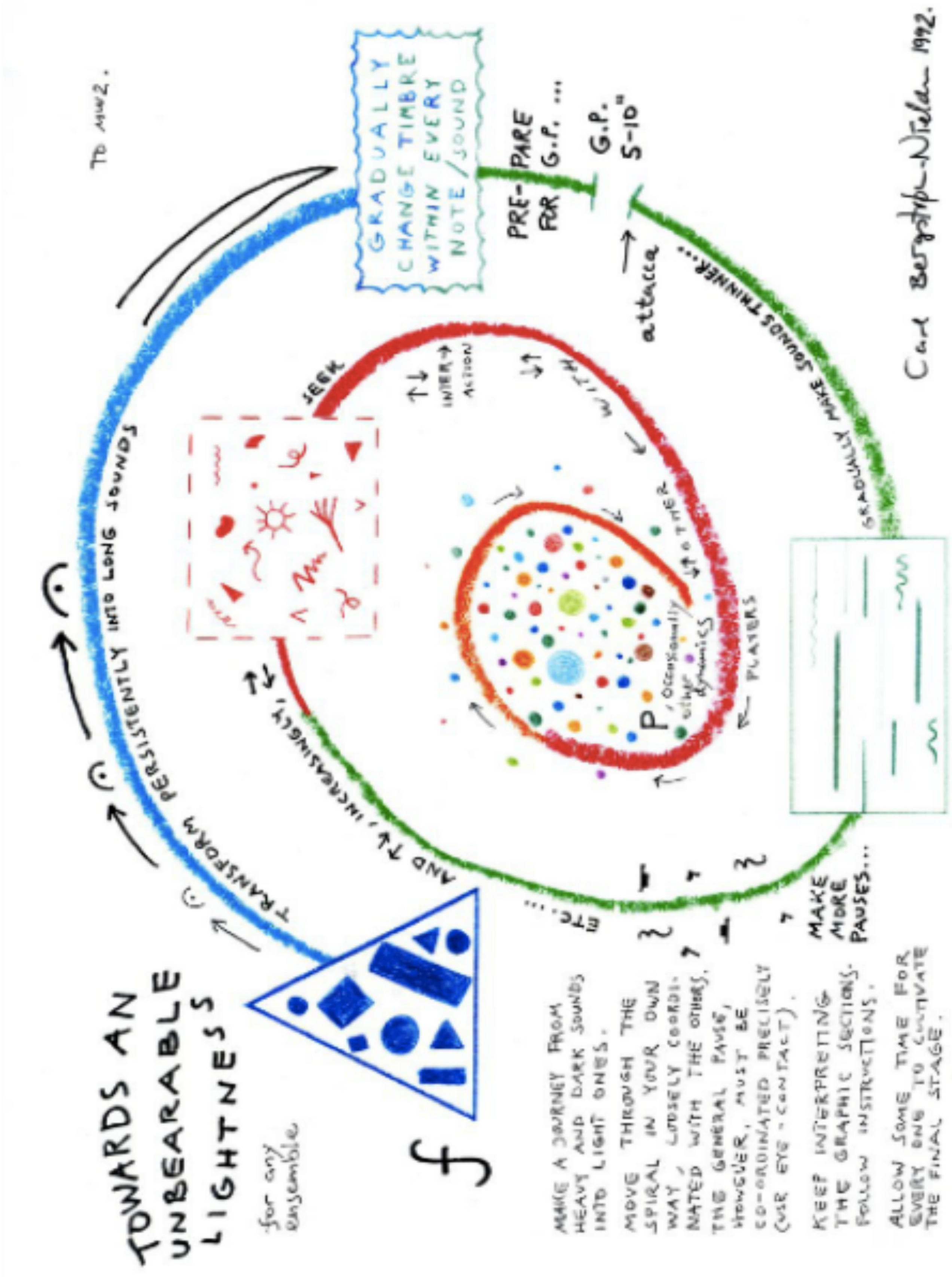

Figure 22: Graphic score by Carl Bergstrom-Neilsen

References

1. Ayad, L.A., Chemillier, M. and Pissis, S.P., 2018. Creating improvisations on chord progressions using suffix trees. *Journal of Mathematics and Music*, *12*(3), pp.233-247.
2. Dannenberg, R.B., 2012. Human computer music performance. In *Dagstuhl Follow-Ups* (Vol. 3). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
3. Siphocly, N.N.J., El-Horbaty, E.S.M. and Salem, A.B.M., 2019. Analysis of Computational Intelligent Techniques of Music Generation. *Egyptian Computer Science Journal*, *43*(3).
4. Cope, D., 2000. *The algorithmic composer* (Vol. 16). AR Editions, Inc.
5. McLean, A. and Dean, R.T. eds., 2018. *The Oxford handbook of algorithmic music*. Oxford University Press.
6. Maurer, J. (1999, March). A Brief History of Algorithmic Composition. Retrieved January 13, 2020, from https://ccrma.stanford.edu/~blackrse/algorithm.html
7. Alpern, A., 1995. Techniques for algorithmic composition of music. *On the web: http://hamp.hampshire.edu/adaF92/algocomp/algocomp*, p.120.
8. Burkholder, J.P., Grout, D.J. and Palisca, C.V., 2019. *A History of Western Music: Tenth International Student Edition*. WW Norton & Company.
9. Roads, C. and Strawn, J., 1996. *The computer music tutorial*. MIT press.
10. Ruttkay, Z., 1997. Composing Mozart variations with dice. *Teaching Statistics*, *19*(1), pp.18-19.
11. Schwartz, E., 1993. *Music since 1945: issues, materials, and literature*. Cengage Learning.
12. Copeland, B.J. and Long, J., 2017. Turing and the history of computer music. In *Philosophical Explorations of the Legacy of Alan Turing* (pp. 189-218). Springer, Cham.
13. Foy, N., 1974. The Word Games of the Night Bird (Interview with Christopher Strachey). *Computing Europe*, *15*, pp.10-11.
14. Assayag, G., Bloch, G., Chemillier, M., Cont, A. and Dubnov, S., 2006, October. Omax brothers: a dynamic topology of agents for improvization

learning. In *Proceedings of the 1st ACM workshop on Audio and music computing multimedia* (pp. 125-132). ACM.

15. Cope, D., 1989. *New directions in music*. William C. Brown.
16. Nika, J. and Chemillier, M., 2012, September. Improtek: integrating harmonic controls into improvisation in the filiation of OMax. In *International computer music conference (ICMC)* (pp. 180-187).
17. Tatar, K. and Pasquier, P., 2019. Musical agents: A typology and state of the art towards Musical Metacreation. *Journal of New Music Research*, *48*(1), pp.56-105.
18. Gifford, T., Knotts, S., McCormack, J., Kalonaris, S., Yee-King, M. and d'Inverno, M., 2018. Computational systems for music improvisation. *Digital Creativity*, *29*(1), pp.19-36.
19. Herremans, D., Chuan, C.H. and Chew, E., 2017. A functional taxonomy of music generation systems. *ACM Computing Surveys (CSUR)*, *50*(5), p.69.
20. Kirke, A. and Miranda, E.R., 2013. An overview of computer systems for expressive music performance. In *Guide to computing for expressive music performance* (pp. 1-47). Springer, London.
21. Edwards, M., 2011. Algorithmic composition: computational thinking in music. *Communications of the ACM*, *54*(7), pp.58-67.
22. Briot, J.P., Hadjeres, G. and Pachet, F.D., 2017. Deep learning techniques for music generation--a survey. *arXiv preprint arXiv:1709.01620*.
23. Bloch, G., Dubnov, S. and Assayag, G., 2008, August. Introducing video features and spectral descriptors in the omax improvisation system. International Computer Music Conference '08.
24. Dubnov, S., Assayag, G., Bejerano, O.L.G. and Lartillot, O., 2003. A system for computer music generation by learning and improvisation in a particular style. *IEEE Computer*, *10*(38).
25. Assayag, G., Bloch, G. and Chemillier, M., 2006, May. Omax-ofon. *Sound and Music Computing (SMC)*
26. Lévy, B., 2009. Visualising omax. *Master II ATIAM. UPMC-IRCAM*.
27. Assayag, G. and Bloch, G., 2007, August. Navigating the oracle: A heuristic approach. I*nternational Computer Music Conference*, vol. 7, 2007, pp. 405–412.

28. Allauzen, Cyril, Maxime Crochemore, and Mathieu Raffinot. 1999. "Factor Oracle: A New Structure for Pattern Matching." In *SOFSEM'99: Theory and Practice of Informatics*, edited by Jan Pavelka, Gerard Tel, and Miroslav Bartošek, 295–310. Berlin: Springer.
29. Bonnasse-Gahot, L., 2014. An update on the SOMax project. *Ircam-STMS, Tech. Rep. Internal report ANR project Sample Orchestrator 2, ANR-10-CORD-0018*, 2014
30. Lévy, B. (2013, December 16). Ressources.ircam. Retrieved January 7, 2020, from https://medias.ircam.fr/xfde422_principles-and-architectures-for-an-intera
31. Lévy, B., Bloch, G. and Assayag, G., 2012, May. OMaxist dialectics. In *Proceedings of the Inter- national Conference on New Interfaces for Musical Expression*. 137–140.
32. The OMax Project Page. (n.d.). Retrieved December 13, 2019, from http://recherche.ircam.fr/equipes/repmus/OMax/
33. Déguernel, K., Nika, J., Vincent, E. and Assayag, G., 2017, July. Generating equivalent chord progressions to enrich guided improvisation: application to Rhythm Changes. in *Proc. of the Sound and Music Computing Conference*, 2017, pp. 399–406
34. Assayag, G. and Dubnov, S., 2004. Using factor oracles for machine improvisation. *Soft Computing*, *8*(9), pp.604-610.
35. Nika, J., Chemillier, M. and Assayag, G., 2016. Improtek: introducing scenarios into human-computer music improvisation. *Computers in Entertainment (CIE)*, *14*(2), p.4.
36. Nika, J., Déguernel, K., Chemla, A., Vincent, E. and Assayag, G., 2017, October. DYCI2 agents: merging the" free"," reactive", and" scenario-based" music generation paradigms. In *Proceedings of the International Computer Music Conference*, Shangai.
37. Bonnasse-Gahot, L., 2010. Donner à OMax le sens du rythme: vers une improvisation plus riche avec la machine. *École des Hautes Études en sciences sociales, Tech. Rep.*
38. Cont, A., 2008, August. ANTESCOFO: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music. In *Proceedings of the International Computer Music Conference*, 2008.

39. Poirot, P.D.G., 1991. SVP: A MODULAR SYSTEM FOR ANALYSIS, PROCESSING AND SYNTHESIS OF SOUND SIGNALS. In *Proceedings of the... International Computer Music Conference* (p. 161). Computer Music Association.
40. Nika, J. (n.d.). ImproteK - H. Sellin playing with B. Holiday, E. Piaf & E. Schwarzkopf. Retrieved January 15, 2020, from https://vimeo.com/146571959
41. Kapoula, Z., Volle, E., Renoult, J. and Andreatta, M. eds., 2018. *Exploring Transdisciplinarity in Art and Sciences*. Springer.
42. Chemillier, M., 2018. De la simulation dans l'approche anthropologique des savoirs relevant de l'oralité. Le cas de la musique traité avec le logiciel Djazz et le cas de la divination. *Transposition. Musique et Sciences Sociales*, (Hors-série 1).
43. Vigne, L.M. and Imbert, R., 2018. De la simulation dans l'approche anthropologique des savoirs relevant de l'oralité: le cas de la musique traité avec le logiciel ImproteK et le cas de la divination. *Transposition. Musique et sciences sociales* Hors série 2. https://journals.openedition.org/transposition/1685
44. Chemillier, M. and Nika, J., 2015. «Étrangement musical»: les jugements de goût de Bernard Lubat à propos du logiciel d'improvisation ImproteK. *Cahiers d'ethnomusicologie. Anciennement Cahiers de musiques traditionnelles*, (28), pp.61-80.
45. Cont, A., Dubnov, S. and Assayag, G., 2006, September. Anticipatory model of musical style imitation using collaborative and competitive reinforcement learning. In *Workshop on Anticipatory Behavior in Adaptive Learning Systems* (pp. 285-306). Springer, Berlin, Heidelberg.
46. Déguernel, K., Vincent, E. and Assayag, G., 2018. Probabilistic factor oracles for multidimensional machine improvisation. *Computer Music Journal*, *42*(02), pp.52-66.
47. Meyer, L.B., 2008. *Emotion and meaning in music*. University of chicago Press.
48. Hofstadter, D.R., 1979. *Gödel, escher, bach*. Hassocks, Sussex: Harvester press.
49. Lewis, G.E., 2018. Why Do We Want Our Computers to Improvise?. *The Oxford Handbook of Algorithmic Music*, p.123.

50. Marcus, G. and Davis, E., 2019. *Rebooting AI: building artificial intelligence we can trust*. Pantheon.
51. Marcus, G. (n.d.). GPT-2 and the Nature of Intelligence. Retrieved February 5, 2020, from https://thegradient.pub/gpt2-and-the-nature-of-intelligence/
52. Somers, J. (2015, August 24). The Man Who Would Teach Machines to Think. Retrieved February 8, 2020, from https://www.theatlantic.com/magazine/archive/2013/11/the-man-who-would-teach-machines-to-think/309529/
53. McVicar, M., Santos-Rodríguez, R., Ni, Y. and De Bie, T., 2014. Automatic chord estimation from audio: A review of the state of the art. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, *22*(2), pp.556-575.
54. Chomsky, N., 2013. *Studies on semantics in generative grammar* (Vol. 107). Walter de Gruyter.
55. Carsault, T., Nika, J. and Esling, P., 2019. Using musical relationships between chord labels in automatic chord extraction tasks. *arXiv preprint arXiv:1911.04973*.
56. Crook, H., 1991. *How to improvise: an approach to practicing improvisation*. Advance Music.
57. Malt, M. and Jourdan, E., 2008, July. Zsa. Descriptors: a library for real-time descriptors analysis. In 5th Sound and Music Computing Conference, p. 134-137, Berlin, Allemagne. See http://www.e--j.com/?page_id=83
58. Liuni, M. and Röbel, A., 2013. Phase vocoder and beyond. *Musica/Tecnologia*, pp.73-89.
59. Průša, Z. and Holighaus, N., 2017, August. Phase vocoder done right. In *2017 25th European Signal Processing Conference (EUSIPCO)* (pp. 976-980). IEEE.
60. Bjerstedt, S., 2014. Storytelling in jazz improvisation. *Implications of a rich intermedial metaphor. Lund: Lund University Publications*.
61. Barrett, F.J., 2000. Cultivating an aesthetic of unfolding: Jazz improvisation as a self-organizing system. *The aesthetics of organization*, pp.228-245.
62. Berliner, P.F., 2009. *Thinking in jazz: The infinite art of improvisation*. University of Chicago Press.

63. http://jazzprofiles.blogspot.com/2015/06/storytelling-in-jazz-composing-in-moment.html
64. Engel, J. and Roberts, A., 2018. Magenta Studio. Retrieved February 02, 2020, from https://www.youtube.com/watch?time_continue=1728&v=akBYPx8MxIE&feature=emb_title, created by Loop.
65. DYCI2: Overview. (2018, June 14). Retrieved January 20, 2020, from http://repmus.ircam.fr/dyci2/project
66. Ficher, T. and Flynn, W. 2016. Jazz Guitar Duets, Mel Bay Publications
67. Srinivasamurthy, A., Holzapfel, A. and Serra, X., 2014. In search of automatic rhythm analysis methods for turkish and indian art music. *Journal of New Music Research*, *43*(1), pp.94-114.
68. Vassilakis, D., Georgaki, A. and Anagnostopoulou, C., (2019). "JAZZ MAPPING" AN ANALYTICAL AND COMPUTATIONAL APPROACH TO JAZZ IMPROVISATION. SMC, MALAGA 2019 PROCEEDINGS.
69. Huron, D.B., 2006. *Sweet anticipation: Music and the psychology of expectation*. MIT press.
70. Greene, T. (2020, January 27). Why the smartest AI is still dumber than a toddler -- and how we can fix that. Retrieved January 29, 2020, from https://thenextweb.com/artificial-intelligence/2020/01/27/why-the-smartest-ai-is-still-dumber-than-a-toddler-and-how-we-can-fix-that/
71. Σπυρίδης, Χ., ΕΥΚΛΕΙΔΟΥ: Κατατομή Κανόνος, Σημειώσεις μαθήματος, Τμήμα Μουσικών Σπουδών Πανεπιστημίου Αθηνών, ΕΚΠΑ.
72. Weinberg, G., Raman, A. and Mallikarjuna, T., 2009, March. Interactive jamming with Shimon: a social robotic musician. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (pp. 233-234).
73. Ghedini, F., Pachet, F. and Roy, P., 2016. Creating music and texts with flow machines. In *Multidisciplinary contributions to the science of creative thinking* (pp. 325-343). Springer, Singapore.
74. Roberts, A.; Engel, J.; Mann, Y.; Gillick, J.; Kayacik, C.; Nrly, S.; Dinculescu, M.; Radebaugh, C.; Hawthorne, C.; and Eck, D. 2019. Magenta studio: Augmenting creativity with deep learning in ableton live. In *Proceedings of the International Workshop on Musical Metacreation (MUME).*

75. Ligon, B., 1996. *Connecting chords with linear harmony*. Hal Leonard Corporation.
76. Coltrane, J. and Coan, C., 1995. *John Coltrane Solos*. Hal Leonard Corp.
77. Goldsen, M.H., 1978. Charlie Parker omnibook: For c instruments. *Atlantic Music Corp*.
78. Stravinsky, I., 1958. *Stravinsky: an autobiography*. Published by M. & J. Steuer.
79. Stravinsky, I., 1964. *Interview,* September 27, New York Times Magazine.
80. Knowles, J.D. and Hewitt, D., 2012. Performance recordivity: Studio music in a live context. In Burgess, Richard James & Isakoff, Katia (Eds.) *7th Art of Record Production Conference*, 2 - 4 December 2011, San Francisco State University, San Francisco, CA.
81. Schechner, R., 2017. *Performance studies: An introduction*. Routledge.