

*In Search of Computer Music Analysis: Music Information Retrieval, Optimization, and Machine Learning
from 2000-2016*

Felicia Nafeeza Persaud

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MA degree in Music

Department of Music
Faculty of Arts
University of Ottawa

© Felicia Nafeeza Persaud, Ottawa, Canada, 2018

Table of Contents

Abstract	vii
Acknowledgements	viii
Glossary	ix
Chapter 1- Introduction and Literature Review	1
1.1.1 General Mission Statement	2
1.1.2 A Critical Overview of Computer Music Analysis: Music Information Retrieval, Optimization, and Machine Learning	3
1.1.3 Persaud’s Five Critical Issues.....	4
1.2 A Sketch of the Relationship between Computers and Music	9
1.2.1 Composition and Performance	9
1.2.2 Applications in Music Theory and Analysis	12
1.2.2.1 Recurrent features: Databases	12
1.2.2.2 Structural Models: Analysis and Counterpoint	14
1.2.3 Music Information Retrieval Versus Optimization	15
1.3 Literature Review	17
1.3.1 David Temperley The Cognition of Basic Musical Structures (2001).....	17
1.3.2 David Temperley and Christopher Bartlette “Parallelism as a Factor in Metrical Analysis” (2002).....	19

1.3.3 David Temperley <i>Music and Probability</i> (2007)	20
1.3.4 David Huron “Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles” (2001).....	21
1.3.5 Darrell Conklin and Ian H. Witten “Multiple Viewpoint Systems for Music Prediction” (1995).....	22
1.4 Conclusion	23
Chapter 2- Music Information Retrieval	25
2.1 Introduction	25
2.1.1 MIR Overview and Applications.....	26
2.2 The MIR Tools	30
2.2.1 Vocalsearch	30
2.2.2 SIMSSA.....	32
2.2.3 Donnelly and Sheppard Bayesian Network Algorithm	37
2.3 Critical Analysis	38
2.3.1 VocalSearch.....	38
2.3.2 SIMSSA.....	39
2.3.3 Bayesian Networks	41
Chapter 3-Optimization	42
3.1 Preference Rules	45
3.1.1 Metrical Structure	45

3.1.2 Contrapuntal Structure.....	49
3.1.3 Tonal-Pitch Class Representation and Harmonic Structure	51
3.1.4 Melodic Phrase Structure.....	53
3.1.5 Parallelism	54
3.2 Probabilistic and Statistical models	55
3.2.1 Introduction	55
3.2.2 David Temperley’s use of Bayesian Probability	57
3.2.3 Statistics and Harmonic Vectors.....	61
3.2.4 Distinctive Patterns using Bioinformatics and Probability.....	63
3.3 Critical Analysis: Optimization	65
3.3.1 Preference rules: Metrical Structure	66
3.3.2 Preference Rules: Counterpoint.....	66
3.3.3 Preference Rules: Tonal-Class representation and Harmony	67
3.3.4 Melodic Phrase Structure and Parallelism.....	68
3.3.5 Probability and Statistics	69
Chapter 4-Machine Learning	71
4.1 Introduction to Machine Learning	71
4.2 Outline of Selected Tools	72
4.2.1 Ornamentation in Jazz Guitar	72
4.2.2 Melodic Analysis with segment classes	73

4.2.3 Chord sequence generation with semiotic patterns	74
4.2.4 Analysis of analysis	75
4.3 Summary	78
Chapter 5- Conclusion	80
5.1 Further Temperley Research and Probability	80
5.2 Machine Learning as a means to an end	81
5.3 CompMusic as an example of Intersection	82
5.4 Five general areas for improvement in the field	83
5.5 Persaud's Five Critical Issues with Solutions	86
Bibliography	89

Table of Figures

Figure 1 Graphic representation of the five critical issues.....	8
Figure 2 Graphic representation of MIR.....	29
Figure 3 Graphic Representation of Optimization	44
Figure 4 Beat Hierarchy.....	46
Figure 5 Graphic of five critical issues with solutions.....	87

Abstract

My thesis aims to critically examine three methods in the current state of Computer Music Analysis. I will concentrate on Music Information Retrieval, Optimization, and Machine Learning. My goal is to describe and critically analyze each method, then examine the intersection of all three. I will start by looking at David Temperley's *The Cognition of Basic Musical Structures* (2001) which offers an outline of major accomplishments before the turn of the 21st century. This outline will provide a method of organization for a large portion of the thesis. I will conclude by explaining the most recent developments in terms of the three methods cited. Following trends in these developments, I can hypothesize the direction of the field.

Acknowledgements

I have appreciated all the help I have had in this thesis writing process. From professors, to friends, to family, everyone deserves a thank you.

Firstly, I must thank my thesis supervisor Dr. P. Murray Dineen who has guided me throughout this process. His feedback and support has helped me immensely to improve as a writer. I am grateful that Dr. Dineen has helped me to gain invaluable skills over the last two years in my Master of Arts.

I would also like to thank my committee members, Dr. Roxanne Prevost and Dr. Jada Watson, who have provided amazing feedback and discussion. They have helped me greatly in creating the final thesis. I would like to thank Dr. Julie Pedneault-Deslauriers as well for serving as a member of the committee for the thesis proposal.

I am grateful to the rest of my professors and colleagues at the University of Ottawa for everything I have learned at the University of Ottawa. It has helped to guide me in creating this thesis and has helped me improve myself.

My friends and family also deserve a thank you for going through sections and drafts throughout this process. A special thank you to my dad, sister and fiancé who went through my first draft. It has come a long way since then.

Glossary

Algorithm: a set of steps followed in calculations or problem-solving operations to achieve some end result.

Computer Music Analysis: analysis of music using a computing software or algorithms. This is a ‘catch all’ term referring to all of the smaller aspects using computers for music analysis including, Music Information Retrieval (MIR), Optimization, and Machine Learning. According to a 2016 book, entitled *Computational Music Analysis*, by David Meredith, a general definition is “using mathematics and computing to advance our understanding of music [...] and how music is understood.” (Meredith 2016)

Machine Learning: teaching of a computer to analyse and find features, so as to gain knowledge of musical conventions. Machine learning is a route that is parallel to MIR, Preference Rule Systems (PRSs), and Probabilistic models. Like a human learning, “a computer learns to perform a task by studying a training set of examples.” (Louridas and Ebert 2016) Following this, a different example is given, and the effectiveness is measured in several ways depending on the task.

Music Information retrieval (MIR): research concerned with making all aspects of a music file (melody, instrumentation, form etc.) searchable. MIR will eventually lead to a search engine for music.

Optimization: a term used in calculus or business that refers to maximizing use of space or resources. Resources are still important in the musical sense, but they refer to time and energy. This is done through accessibility and more efficient computer tools and algorithms. Examples given below to show that it is possible to optimize analysis by integrating more mathematics and computer tools.

Piano-roll input: a graphic representation of a score with notes on the vertical axis and timing in milliseconds on the horizontal.

Preference rule system (PRS): a set of instructions for a computer in a hierarchy. These can be created as a system where there are multiple sets with a hierarchy within “criteria for evaluating possible analysis of a piece.” (Preface Temperley and Bartlette 2002) This is known as a rule-based grammar in Manning et al 2001.

Parallelism rule (as a type of preference rule): the idea that the similar construction of a musical element be regarded as important in a PRS.” Prefer beat intervals of a certain distance to the extent that repetition occurs at that distance in the vicinity.” (Temperley and Bartlette 2002, 134)

Probabilistic Methods: a method of analysis based in probability. The word “Probabilistic” means for an idea to be based on or adapted to a theory of probability, this term encompasses even distant uses of probability in computer models. This is a term used by Temperley referring to a computational method that uses probability.

Chapter 1- Introduction and Literature Review

1.1 Overview

My interest in Computer Music Analysis stems from my fascination with interdisciplinarity in music analysis. Computer Music Analysis intersects with mathematics, computer science, psychology, and, of course, music. My thesis will take a small sampling of interdisciplinary tools in Computer Music Analysis from Music Information Retrieval (MIR), Optimization, and Machine Learning. MIR aims to make music searchable, primarily through online databases. Optimization encompasses many different tools with the eventual goal to understand human perception of music. Machine Learning, on the other hand, teaches the machine, often a computer, to perform a task, making the tool itself the end goal.

For this thesis, I preface my work with Peter Manning's entry, entitled "Computers and Music," in the Grove Dictionary of Music and Musicians as a way to understand the existing conventions and uses of computers in music prior to the year 2000. Manning does not offer a specific definition, but instead discusses the common uses and devices of the computer as it relates to music. He states, "Computers have been used for all manner of applications, from the synthesis of new sounds and the analysis of music in notated form to desktop music publishing and studies in music psychology; from analysing the ways in which we respond to musical stimuli to the processes of music performance itself." (Manning et al 2001). This quote exemplifies how interdisciplinary Computer Music Analysis is. Manning's work touches on composition, performance, and analysis addressing a key critical issue: human error. A computer is only useful because of its human programmer no matter what the application. With every new application of the computer—or tool—there are more issues and limitations. For example, a tool

that identifies duple metrical structures cannot identify compound meter and has a margin of error. The idea of the human creation of a computer model, and its limitations, is the focus of my thesis and is explored in three branches of Computer Music Analysis: Music Information Retrieval (MIR), Optimization, and Machine Learning. Manning's entry coupled with the Literature Review provide a foundation on which I build this thesis.

1.1.1 General Mission Statement

This thesis aims to critically examine specific tools in Music Information Retrieval (MIR), Optimization—a term referring to improvements in Preference Rule Systems and Probabilistic Models—and Machine Learning individually. The exploration of MIR, Optimization, and Machine Learning will do two things: act as a survey of the literature and show trends within these subfields. In the conclusion, I show how the three aspects can interact. Most branches in Computer Music Analysis run in parallel (Meredith 2016), and few researchers take inspiration from the parallel branches. It is not my intent to show that there is *no* interaction, but merely to show opportunities for more interaction.

To survey the literature, I first look at the developments—prior to the turn of the 21st century, the period when the field of Computer Music Analysis was born. The background comes primarily from David Temperley's book the *Cognition of Basic Musical Structures* (2001) as well as from works covered in the Literature Review and the Sketch of the Computer-Music Relationship sections. To explore current trends, I restrict myself primarily to the literature from 2000 to 2016. These texts build from the turn of the century and show how researchers utilize new technology to push the field further. This area constitutes the body of the thesis and shows

where the field has gone and where it is going. Additionally, using a critical examination of the literature, I explore recent trends in Computer Music Analysis and offer points of entry for new research. I use models drawn from World Music Research. I concentrate on the three areas of the field, as mentioned above. This research can be applied to other similar areas like Mathematical Music Theory, which represents basic musical structures in a mathematical form, or Computational Musicology, which investigates the simulation of computer models in music.

1.1.2 A Critical Overview of Computer Music Analysis: Music Information Retrieval, Optimization, and Machine Learning

The current state of the field in Computer Music Analysis sees a shifting of positions among the three areas: Music Information Retrieval (MIR), Machine Learning, and Optimization. Music Information Retrieval is the most rapidly evolving field of the three; due in large part to developments in and the spread of computers and the Internet – specifically an increase in computing capacity. The second field is Machine Learning; this is similarly due to computing capacity and the Internet, but also because of its widespread use in other disciplines, which music researchers are drawing from at greater and greater lengths. The third field is Optimization, which has stagnated. However, Optimization borrows from other disciplines, and contributes to the advances made by MIR and Machine Learning. As such, we can see that Optimization is currently evolving, even if other two fields are moving at a much greater pace.

To sketch in greater detail, there are crucial differences and overlapping areas between the three fields that explain their current situations. Machine Learning is a precise endeavor that aims to create specific tools to meet well-defined goals or serve finite tasks. MIR, on the other hand, works with large bodies of data and serves goals that are often ill-defined if not undefined.

Conversely, Optimization is presently in a state of coming together— in fields other than music — and, therefore, would appear not to be advancing as quickly. But, in fact, Optimization in its current state is laying a framework for major developments.

Though there is overlap between MIR, Optimization, and Machine Learning, it is limited to a few researchers and projects. Examples include the following: Darrell Conklin using probability and bioinformatics in conjunction with Machine Learning; Giraud et al, who are creating a tool for MIR and Optimization; and, most notably, CompMusic—a database for six subsets of World Music—that uses both Optimization and Machine Learning to create an MIR database. These will be discussed in the later parts of the thesis.

1.1.3 Persaud’s Five Critical Issues

From the critical perspective adopted in this thesis, several issues arise. Some of which have been addressed in the literature surveyed. Unfortunately, they have not been brought together in such a fashion to yield an overall critical perspective of the current field. To this end, I have isolated five central critical issues, which I address here. During the remainder of the thesis, I make reference to these from time to time, by means of a numbered list set out below and in Fig. 1. I refer to these as Persaud Critical Issues, since, to my knowledge they have not been catalogued in this fashion.

Persaud’s Critical Issue 1. Human Error.

Firstly, data entry is still largely human-dependent and with large amounts of data—like with an MIR database—a person will often make mistakes. This was discussed by both Peter Manning in his definition and David Huron about *The Humdrum Toolkit*. As Huron and Manning

explain, the machine is limited by the programmers themselves. Outside of research, artificial intelligence (AI) is being used to complete simple tasks and can learn, by itself, various other tasks. Similarly, quantum computers are becoming more common instead of using simple binary code. Both of these devices are making their way into day-to-day life and eventually will end up in multidisciplinary research. In terms of what is being used currently, data entry could be improved by the application of Machine Learning. Certain parameters could be handled by machine input rather than human input. These advances are being made elsewhere but have not been seen in the area of music research, except in world music database creation [see conclusion of the thesis]. We need to see more inroads made by Machine Learning in the analysis of Western music and Ancient music.

Human limitations are not only evident in data entry but also in setting parameters, in annotations, and in the creation of algorithms in general (Huron 1988). Setting parameters is a vital aspect of Optimization. It enables the most accurate analysis of the data provided and, therefore, generate more accurate outcomes. Because the parameters are calibrated by humans, there is an implicit limitation. This similar to the annotation of pieces in MIR databases and the creation and application of an algorithm in Machine Learning.

Persaud's Critical Issue 2. Input Specification

Input modes are not well-defined by researchers to be easily understood. To a certain extent, this is a problem of writing and communication, one that arises from research silos. This could be resolved by creating common standards and modes of discourse for describing computer research in music, and specifically the modes of input involved. Complementary to input specification due to research silos, input modes change from generic type to type. For

example, popular music is not often scored, while ancient music is not performed in its original form. As such, the input for popular music would most likely be an audio file, while for ancient music, an image of a score is more likely. Furthermore, the input could differ from a full form, such as all tracks on a song, to a simpler form, such as main melody only. This further complicates the situation.

In addition to genre, input modes depend upon translation into computer compatible formats. Though an MP3 audio format is widely available, it is not easily readable for analytical us. As a work around, researchers use either a MIDI format, or the input is further broken down into tracks. In the study of ancient music, image data cannot be read by a computer and must endure multiple passes of analysis using computer-based algorithms and processes, but this method still yields errors.

Persaud's Critical Issue 3. No Consistent Mode of Evaluation for Non-MIR Tools

Music Information Retrieval Evaluation eXchange (MIREX), is a method of formally evaluating MIR systems and algorithms. This does not exist for other branches of Computer Music Analysis like Optimization and Machine Learning. These unknown standards for algorithms and tools result in an end-product that may not have any further use beyond its creation. Furthermore, without a widespread knowledge of the tools and algorithms, they cannot be used for MIR or other branches of Computer Music Analysis simply due to unknowingness.

Persaud's Critical Issue 4. The Interdisciplinary Problem (Downie 2003)

The Interdisciplinary Problem is one that is examined and discussed by Stephen J. Downie in his article "Music Information Retrieval." Though this is an issue in MIR specifically,

it extends to other branches of Computer Music Analysis such as Optimization and Machine Learning. It simply refers to the lack of coordination between researchers and research fields when it comes to creating a tool and the different uses of the same terminology. Some tools and systems are made overly difficult for someone without programming knowledge, even though the outcomes of the tool would be useful to them.

Persaud's Critical Issue 5. "What's the point?" Lack of Defined Goals and Frameworks

Research in Computer Music Analysis often comes as small creations and discoveries rather than a large finished tool. As Computer Music Analysis often concentrates on the method to an output, these smaller steps cannot be used by another researcher until it is completed. Furthermore, the specific usage of the individual step is unknown or has very few applications, if any, so the "What's the point?" argument returns. This argument also does not take into account the full potential of each field and is created by a lack of understanding for the goals of each branch in Computer Music Analysis.

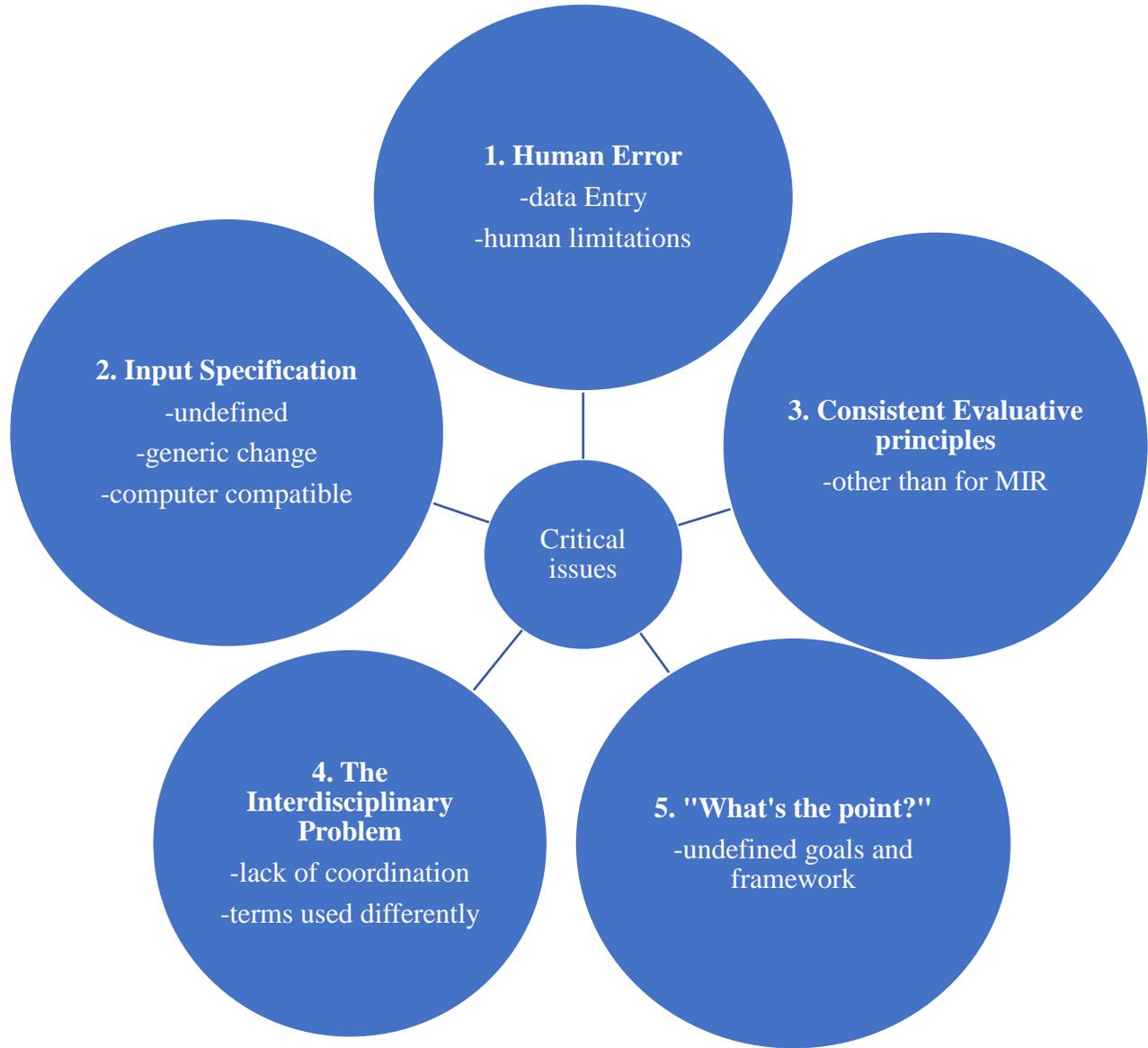


Figure 1 Graphic representation of the five critical issues

1.2 A Sketch of the Relationship between Computers and Music

1.2.1 Composition and Performance

Music and computers have a lengthy history that touches on three fields: composition, performance, and music research. To understand the current state of Computer Music Analysis, the history needs to be discussed. In fundamental terms, the above-mentioned disciplines helped shape Computer Music Analysis

In terms of composition, computer music was one of the principal areas of early research. One main source for understanding this research was the *Computer Music Journal*, founded in 1977. This journal examines crossroads between computers and music such as composition with computers, MIDI, synthesizer theory, and analytical models using the computer (Computer Music Journal). Though the material is broad, there have been specific issues that address analytical models included in this thesis. This publication includes articles about CompMusic—an organization committed to database creation for World Music—which I will return to in my conclusion. The publication also includes Donnelly and Sheppard’s “Classification of Timbre Using Bayesian Networks” which is one of the few instances of cross-branch research.

While the original inroads made into computer music composition were slow and burdened by clumsy and awkward hardware, this situation soon changed. Curtis Roads is a composer of electronic music and an author. His 1985 book, *Composers and the Computer*, is interview-based to get the composer’s perspective. According to Appleton’s review, Roads’s main point is that arts and science are becoming closer to create new music (Appleton 1986). Furthermore, Appleton explains the importance of understanding the means in music creation and the method of computer usage is vital for listening to computer music compositions. “If [...]

the principles of serial technique are necessary to an intelligent hearing of the works of Webern, Carter, Babbitt, or Boulez, then surely an appreciation of the principles of algorithmic compositional techniques and the possibilities of digital sound synthesis are required for the through audition of works by Xenakis, Chowning, Risset, and Dodge (Appleton 1986, 124).” This quote situates the importance of method in music and how the new computer capabilities enhance the composition process.

In 1986, a symposium on computer music composition was held and a review was written in the *Computer Music Journal*. This symposium was a “product of a questionnaire sent in 1982, 1983, and 1984, to over 30 composers experienced in the computer medium” (Roads et al, 40). The review examines, in a similar manner to Roads’ book, what brought the composer to the computer and how they choose to use it. The review states that “articles in *Computer Music Journal* and other publications point to the broad application of computers in musical tasks, especially to sound synthesis, live performance, and algorithmic or procedural composition” (Roads et al 1986, 40).

Music Representation Languages (MRLs) are another important milestone in the history of Computer Music Analysis. An MRL is a type of format that the computer can understand (Downie 2003), and these are vital to composition. An example of this is Musical Instrument Digital Interface commonly known as MIDI. MIDI revolutionized sound processing by enabling the user to store real input, such as playing on a synthesizer, into movable and changeable blocks of sound easily understood by the computer. It has two-way variability because there is a disparity from the player of an external synthesizer and the producer can move and change the blocks of sound after the player has played (Manning et al 2001). It provides more control to all parties for its end result and MIDI is now widely used.

Another significant creation in computer music composition is music notation software. This software, like Finale, often include a MIDI playback. According to Manning “it quickly became apparent that major composition and performance possibilities could be opened up by extending MIDI control facilities to personal computers” (Manning et al 2001, 169). This new MIDI playback on music notation software gave the composer the ability to create music digitally with the option to hear what it would sound like.

Computer music composition, of course, continues today. Recent developments include ChuckK, a programming language specifically for music and is prevalent for laptop orchestra use (Wang et al 2015), and melodic idea generation and evaluation—which is the creation of a motive and the assessment of it (Ponce de Leon et al 2016). Both tools are used for the creation of musical ideas. ChuckK, for example, can create a complete piece in real time. Though computer music composition is important to the relationship between computers and music, it will not be further discussed in this thesis. The field of Computer Music Analysis has moved away sufficiently to be treated as a separate endeavour, at this point.

It should be noted that composition with computers is only one aspect of computer assisted musical creation. According to Manning’s “Computers and Music”, the uses of computers in music can be separated into two branches: performance and music theory. For performance, MIDI is highlighted as a major development, but more performer-like methods are being developed such as DARMS (Manning et al 2001). DARMS is a “comprehensive coding system [...] which has the capacity to handle almost every conceivable score detail” (Manning et al 2001, 176). For current performances, Laptop Orchestra is becoming more prevalent at universities. Though computer use in performance is important, I will not be concentrating on it.

1.2.2 Applications in Music Theory and Analysis

Music research uses for computers are more complex and have been based around two facets:

1. The first is identification of recurrent features. Recurrent features are an important aspect of analysis as it can show that a set of items is a pattern rather than a coincidence. “One of the earliest uses of the computer as a tool for analysis [...] involves the identification of recurrent features that can usually be subjected to statistical analysis.” (Manning et al 2001, 174). Statistical analysis further strengthens a pattern by utilizing quantitative measures. Statistical analysis is still present today and will be discussed in Chapter 3.
2. The second concerns the application of two kinds of “Rule-based analysis.” Analysis used for generative purposes and analysis used in and of itself or as an analytic method. As Manning describes rule-based analysis in general: “rule-based analysis methods presuppose that the processes of composition are bound by underlying structural principles that can be described in algorithmic terms. [...] At this level it becomes possible to establish links with computer-based research into musical meaning” (Manning et al, 174).

Now I will present examples of both facets. Both show, in a simple fashion, the above two ideas and, also, demonstrate the main sources of error and limitation in Computer Music Analysis.

1.2.2.1 Recurrent features: Databases

A major database software for computer music research was the *Humdrum Toolkit* created by David Huron and its files finished revision in 2001. Huron is based at the Ohio State University School of Music and commonly researches Music Cognition, Computational

Musicology, and Systematic musicology. *The Humdrum Toolkit* runs using UNIX software tools, but it is compatible with previous versions of Windows and Mac platforms. This database gives the public access to information on scores, and renotes scores in a format that is useable with the *Humdrum Toolkit*. It is also possible to import or export files from *Finale* software for scores that are not available in the database. “Humdrum” itself is composed of the *Humdrum Syntax* and *Humdrum Toolkit*. The syntax, like other programming language, enables the user to search for files and other elements using the *Humdrum Toolkit*. This programming language, however, must be learned to adequately use the software.

The *Humdrum Toolkit* is used for recurrent features because of its capabilities. The capabilities of Humdrum include searching between sets of pieces for motives, syncopation, harmonic progression, dynamics, pitch, and meter. These elements of music can be searched by genre, by composer, and by any other grouping for an overarching and statistical analysis, therefore, this use for computers in music aligns with Manning’s definition in Grove. However, some of the above-mentioned elements are more easily found using the *Humdrum Toolkit* software than others. Firstly, this is due to “the interdisciplinary problem” since some queries need a complex search using programming language. Programming knowledge is something that is not consistent between all database users. Secondly, human error is always a possibility with a completely manmade database. Like all tools, this one is imperfect. Huron found three reasons for mistakes when using computers because of “Humdrum” (Huron 1988). They are as follows:

1. Errors in actual score
2. Errors in transcription of score
3. Errors by program

These errors according to Huron, are human. ¹

1.2.2.2 Structural Models: Analysis and Counterpoint

P. Howard Patrick in 1978 used computers for analysis of suspensions in the Masses of Josquin Des Prez. Patrick made an important distinction between music theory for the composition student and music theory for the computer rule-based structural model: music theory is often a description, but a computer needs a set of steps to follow. To get the computer to properly parse and identify the data, Patrick looked at the errors and changed criteria as needed. (Alphonse 1988)

Arthur Mendel inspired Patrick's study in a seminar by looking for the criteria of structure in Josquin's work. Patrick outlined the goal of this project as getting computer programs to print a reduction of a score by, first, going through a succession of tests and then finding the "most consonant pitch" (Patrick 1974, 325). Patrick tested three randomly selected texts to outline the problems that he described as "Non-Suspensions (Patrick 1974,326)" and "Problem Suspensions." (Patrick 1974, 328) These errors were due to the computer's now 'preconceived' notion of what a suspension is, but the largest error, as explained by Patrick, are the questions that people ask the computer.

Criticism for this type of analysis is that it only yields a result that can be found by a person doing the research by hand and thus is susceptible to the same kinds of errors humans might make. As stated by Patrick, "The limitations of the computer are overshadowed by the inherent limitations of the user." (Patrick 1974,321) This means that the computer can find any

¹ These sources of error are paraphrased from Huron 1988, 254

solution, but only if it can be fathomed by the user. Some larger scale problems are too difficult to solve without help from another source, such as a computer. In this sense, Patrick thought the computer-aided analysis route was the most useful. This set the groundwork for development in Computer Music Analysis that do not mimic “research by hand.”

1.2.3 Music Information Retrieval Versus Optimization

Music Information Retrieval (MIR) is interdisciplinary, due to its computer-based information, and originated from the same point as Optimization. But, the two fields have different goals.

By *Music Information Retrieval*, I mean the sector of Computer Music Analysis that aims to create a database, either analytical or non-analytical, drawn from characteristics of a musical document such as a score, so as to further research. MIR aims to look into musical documents to find features or commonalities between different works of music. MIR approaches recurrent features by creating a database with annotations, or another searchable method, so a user can search for a specific feature.

Optimization, which concerns itself with preference rules, probability, and statistical models, does not detach itself from the human experience. The following quotation demonstrates the distinctiveness of Optimization for MIR: “Computational research in music cognition tends to focus on models of the human mind, whereas MIR prefers the best-performing models regardless of their cognitive plausibility” (Burgoyne et al 2016, 214). In summary, Optimization is tied to music cognition (Burgoyne et al 2016) while MIR is not.

MIR has turned into an ever-growing and prevalent field due to the internet (Fujinaga and Weiss) and is present in commonly used items like Google Books (Helsen et al 2014), but it

originally came from a small field of research in comparison. According to Burgoyne et al, in 1907, C.S. Myers studied Western folksong using MIR, which required tabulation done by hand examining the intervals present in folksongs. Similarly, in ethnomusicology a year earlier, 1906, a similar method was used to find features in Non-Western music to differentiate it from Western music (Burgoyne et al 2016). The practice of “Finding Features” has become a standard use for Computer Music Analysis. These are the earliest examples of Music Information Retrieval even though the term itself was not used until the 1960s.

From 1907 to the 1960s Music Information Retrieval was ignored, but, “interest grew in computerized analysis of music” (Burgoyne et al, 215) because of the prevalence and accessibility of computers. The beginning of MIR concentrated on methods to input music into the computer (Burgoyne et al) such as notational software or standardized audio file formats like MP3 and MIDI (Fujinaga and Weiss). This made it possible for the computer to ‘understand’ the musical items. These methods grew into more complex software applications like Humdrum which was discussed in section 1.2.2.1

This history of MIR is written in brief, however it gives a basic outline of its developments that is important to the thesis. Since, this field re-emerged because of the internet and increasing availability of computers, the tabulations could be done using a software instead of by hand. After creating a form of music that can be understood by a computer, databases, like Humdrum, were more easily produced. Creating a database of music recognizable by a computer, according to Andrew Hankinson—a Digital Humanities and Medieval Music researcher—, is the first step in a large retrieval system (Helsen et al 2014). Large databases of different varieties will be further discussed Chapter 2.

1.3 Literature Review

I aim to explore the major works I use for this thesis in the literature review. The order is to mirror the order of the thesis: first Optimization then Machine Learning. MIR has a more complicated Literature base, so I discuss it in Chapter 2. I commence with David Temperley's works in chronological order because I incorporate their organization tools and major ideas into Chapter 3. Parallelism is highlighted because it grows from a single-line preference rule to a multi-level set of ideas. Since perception is key to Optimization, I include David Huron for the link from computers to perception. Huron's paper examines voice-leading rules, which are common knowledge and vital to music theorists, thus act as a stable starting point. The final work is Darrell Conklin and Ian Whitten's paper investigating the multiple-viewpoint system. This article is one of the first that examine Machine Learning in music and should, therefore, be included.

1.3.1 David Temperley *The Cognition of Basic Musical Structures* (2001)

David Temperley is centred at the Eastman School of Music and writes extensively on music theory and music cognition. I will concentrate on specific sections of his book *The Cognition of Basic Musical Structures* (2001), that explain Preference Rule Systems or Computational models. Temperley outlines the following six Preference Rule Systems in the first half of the book, Metrical Structure, Melodic Phrase Structure, Contrapuntal Structure, Tonal-Pitch-Class Representation, Harmonic Structure, and Key Structure, and the second half explores the expectation of the listener, Rock Music, African music, composition, and recomposition. The first half of the book is where I will concentrate this review. Temperley states that the goal of the

book is to explore the “‘infrastructural’ levels of music,” meaning the basic building blocks of music perception, because there is very little research on the subject.

Before presenting the Preference Rule System (PRS), Temperley outlines previous research on musical structure as it relates to each section. For example, Temperley describes at length the Desain and Honing model for beat induction in the chapter on Metrical Structure. The specificities of each section is discussed in Chapter 3 of this thesis . He notes that each PRS is based on a piano-roll input for the computer. The PRS itself is a group of rules the computer follows to narrow a set of possible choices. Within each rule there is a preference—hence the name preference rule. The end choice is selected because more rules are preferred in a specific hierarchy.

After presenting Preference Rule Systems, Temperley describes the tests he goes through to ensure well-functioning systems. Meter, unlike the others, has had plenty of research concerning theoretical and computational models. Temperley builds upon the Lerdahl and Jackendoff *Generative Theory of Tonal Music* (1983) by adapting it for a preference rule approach. The meter section takes the *Well Formedness* definition from Lerdahl and Jackendoff where grouping and hierarchy are most important and Temperley explains it as “every event onset must be marked by a beat [and] that a beat at one level must be at all lower levels” (Temperley 2001,30). This is used in all successive PRSs. Similarly, for Key Structure there is sufficient research from music cognition and computational methods to improve upon. Temperley uses the Krumhansl-Schumckler Key-Finding Algorithm and discusses problems and solutions.

The other four PRSs take a list of rules and within each have a list of preferences in a specific order, so the computer knows which item is the most important or most common. For

example, the Phrase Structure Preference Rules (Temperley 2001 Melodic Phrase Structure Chapter pp. 68-70) comprise of three rules.

1. Gap Rule: Prefer to Locate phrase boundaries at
 - a. Large inter-onset intervals and
 - b. Large offset-to-onset intervals
2. Phrase Length Rule: Prefer phrases to have roughly 8 notes
3. Metrical Parallelism Rule: Prefer to begin successive groups at parallel points in the metrical structure

This is for only well-formed, by the previously mentioned definition, monophonic melodies. For implementation of each of these rules, a formula, score or other quantification is applied. The best “score” is the best analysis for a melody.

Temperley’s Preference Rule Systems gives me multiple examples of how the computer evaluated different problems which I can then relate to other models for evaluation. In this regard, Temperley’s 2001 book acts as a springboard for my thesis. It gives important background information in Computer Music Analysis and shows me how Temperley’s subsequent work has built upon it. The book will be further discussed in Chapter 3: Optimization.

1.3.2 David Temperley and Christopher Bartlette “Parallelism as a Factor in Metrical Analysis” (2002)

This text builds upon the previous Temperley book by adding further information to the “Metrical Parallelism Rule.” (Temperley 2001,70). The “well-formedness rule,” as mentioned in Temperley 2001, still applies in this article, as does the need for monophony. The goal of this

article is to build upon the book for clarity, accuracy and precision when dealing with Parallelism.

Temperley and Bartlette examine the effect of Parallelism and realized that the definition must be modified. Parallelism is defined as a repetition either of the exact sequence or the contour. The Parallelism Rule is now redefined to “prefer beat intervals of a certain distance to the extent that repetition occurs at that distance in the vicinity.” (Temperley and Bartlette 2002, 134) This is useful to the thesis because it gives a more inclusive definition to Paralellism as a term and a rule and, also, because of the influence it had on the later treatment of parallelism.

1.3.3 David Temperley *Music and Probability* (2007)

Though Temperley was content with the 2001 book, it seemed like more should be added to the approach because preference rule models could not be applied to “linguistics or vision” (Temperley 2007, ix). The goal of the 2007 book is to use specific Bayesian probability tool, as a link between perception and style. In the perception of linguistics and vision, Bayesian probability techniques such as probability of an event following another are more common in computer analytic tools. To quote Temperley, “I realized that Bayesian models provided the answer to my problems with preference rule models. In fact, preference rule models were very similar to Bayesian models” (Temperley 2007,x) meaning that the existing PRSs can be easily turned into Bayesian models.

The book shows a new trend in Computer Music research: probability. It uses the Essen Corpus, also known as the Essen Folksong Collection,² to test for the central distribution of the

² The Essen Folksong collection is a set of folksongs from Germany, China, France, Russia and more collected by Helmut Schaffrath. <http://essen.themefinder.org/>

aspects of music (and relies on a method of representation created by Lerdahl and Jackendoff in 1983, which, by this point, was familiar to music theorist). The book itself touches on Rhythm, Pitch, Key, Style, Composition, and, like the first computer music analytic tools, error detection in its main chapters.

1.3.4 David Huron “Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles” (2001)

I have included this work in the literature review because we must remember that all computer models tie back to perception, in some way, to be correct. It should be noted that Huron’s text was also referenced in Temperley’s work because the psychological principles behind musical aspects make computational modelling difficult.

Huron’s 2001 work shows the relationship between voice-leading and auditory perception using perception. The article presents a set of the voice-leading rules, then derives them from the perception principles, and finally it makes ties to genre. Each voice leading rules is scrutinized under three questions:

1. What goal is served by the following rule?
2. Is the goal worthwhile?
3. Is the rule an effective way of achieving the purported goal? (Huron 2001, 1)

Huron brings up the important concept of culture. With analysis, it remains unknown if these principles of auditory perception are inherent in all people or if they are created by cultures. However, Huron notes that “perceptual principles can be used to account for a number of aspects of musical organization, at least with respect to Western music” (Huron 2001,1) and

concludes that six principles in perception account for most voice leading rules in Western Music.

Another important aspect brought up is the compositional goals because the composer plays with the perception of the listener. For example, Huron mentions “Bach gradually changes his compositional strategy. For works employing just two parts, Bach endeavors to keep the parts active (few rests of short duration) and to boost the textural density through pseudo-polyphonic writing. For works having four or more nominal voices, Bach reverses this strategy” (Huron 2001, 47). This deceives the listener because a four-voice work may sound more sparse while a two-voice work sounds more dense making these voice-leading rules more like compositional options.

1.3.5 Darrell Conklin and Ian H. Witten “Multiple Viewpoint Systems for Music Prediction” (1995)

Darrell Conklin concentrates on research in Machine Learning and Music at the University of Basque Country in Spain. This article has been cited in Temperley’s works such as *The Cognition of Basic Musical Structures* (2001). The paper takes an “empirical induction approach to generative theory” (Conklin and Whitten 1995, 52) by exploring previous compositions for style and patterns. More specifically, this article uses Bach Chorale as a starting point for choral music.

Conklin and Whitten describe Machine Learning, applied to music research, as follows: “Machine learning is concerned with improving performance as a specific task. Here the task is music prediction” (Conklin and Whitten 1995, 55). Since much of Machine Learning uses context models, but that requires exact matches. Music does not always use exact matches

because *similarity* is enough for auditory perception, Conklin and Witten take a multiple-viewpoint system. Each viewpoint is an aspect of music, to derive musical ideas that take style into account.

Conklin and Whitten describe the next steps in this field as:

1. Research on prediction and entropy of music
2. The creation of “a general-purpose machine learning tool for music” (Conklin and Whitten 1995,71) for all musical genres

Their work adds to the thesis by providing the beginning of Machine Learning. From this, the rest of the accomplishments in Machine Learning and music can be put into perspective.

1.4 Conclusion

In the introductory chapter of this thesis, I have described my goal: to critically examine aspects of Music Information Retrieval (MIR), Optimization, and Machine Learning. Between MIR and Optimization there is a common starting point, but they differ in goal. MIR aims to create a database or multiple databases for further analysis while Optimization uses a computer model to understand the human perception of a musical structure. Machine Learning is different than the other two since it concentrates on the creation of a tool and not necessarily the uses.

I have surveyed specific literature in the field of computer music analysis for a background and inroad to the research from 2000 to 2016. For a historical context, I have brought in Manning’s multi-faceted explanation of the relationship between computers and music. This mentions composition, performance, and analysis and displays the many important developments prior to the turn of the century. The developments include Music Representation Languages (MRLs)—like MIDI—and notation software because they created a widespread

usage. This literature touches on MIR, Optimization, and Machine Learning and, also, exposes some critical issues in Computer Musical Analysis.

I have set out a list of five critical issues, that I use to gain critical perspective on the field. The first issue is Human Error which refers to human limitations and the capacity to make mistakes. This was brought up by both Peter Manning and David Huron. Second is input specification, which is a recurring issue since articles do not specify what input is used for a tool. The input is largely genre-based due to availability. Consistent Evaluative Principles are needed for all branches of Computer Music Analysis, so that there is a reliable set of algorithms and methods to be drawn upon. The Interdisciplinary Problem is an issue with term usage and level differences in tools creation and is evident through all of the authors in the literature review. This is because each author uses their own set of terms based on their usual field of research. “What’s the Point?” refers to the lack of reason for a specific tool because, for a branch like Optimization, the tools are working towards understanding human perception. This means a specific tool may not have a specific usage at its inception. Using this chapter as a basis, I begin my analysis of specific tools in each of the three subfields starting with Music Information Retrieval.

Chapter 2- Music Information Retrieval

2.1 Introduction

Music Information Retrieval (MIR) is a subsection of Computer Music Analysis that is growing exponentially because of current technology. MIR is concerned with examining music, either by locating or by analysing, and often aims to make music searchable. The locating branch is often aimed at examining the metadata of a large set of works. The analysis/production branch concerns itself with a smaller number of pieces but goes into much greater detail (Downie 2003) as is stated by Downie: “Analytic/Production systems usually contain the most complete representation of music information” (Downie 2003, 308). Databases created for MIR can be accessible through the internet, so they are used by all researchers if they have the background knowledge needed.

The goal of this chapter is to begin a critical comparison of tools and problem-solving methods in MIR. This will be accomplished by discussing three projects: a large completed tool, a large tool in progress, and a small tool. These tools are just the “tip of the iceberg” when it comes to MIR, but they have been chosen to show different stages within the evolution of a tool. The large completed tool is VocalSearch where song lyrics can be searched to identify their presence in a song. The in-progress tool is a research project called the Single Interface for Music Score Searching and Analysis (SIMSSA). The small milestone studied here is Patrick Donnelly and John Sheppard’s approach to timbre identification using probability. In fact, Donnelly and Sheppard’s project provide a solution to a specific problem which in turn can provide help to a larger database. This final milestone will show how smaller projects in Computer Music Analysis can help solve larger problems and thus help move the field forward.

2.1.1 MIR Overview and Applications

The purpose of this section is to give a description of major terms in Music Information Retrieval (MIR) and to show the different systems at work in MIR. I will not be going in depth about all systems, but I would like to show the complexity of MIR. I will first explain the two main types of MIR systems: locating and analytic/production. Then I will outline the different types of data. I will then explain how the different types of musical information fit into each of the data categories and systems.

MIR examines multiple facts of music information in many different forms. According to J. Stephen Downie—the creator of MIREX and specializing in information sciences at the University of Illinois—there are two different types of MIR systems: locating and analytic/production (Downie 2003) as mentioned in the introduction.

The locating systems are used by people searching for music either as a consumer on a website or as a researcher in a recordings database. A locating system looks at many works, but does not go in depth, and often locates information on the title, composer, performer etc. This type of information is called metadata. An analytic/production system generally looks at a small number of works, but in much greater detail. These systems, for example, can look at audio recordings, pictures of scores, and/or symbolic forms of scores. (I will not go into detail about specific systems at this point since they will be discussed later in the chapter.)

The different types of possible data in music, as mentioned above, are metadata, audio, symbolic, and image. Metadata is simply data about data, so, in music, this is information about the performers or pieces performed, such as title, composer, etc. Audio data is a recording. Most commonly, MP3 files are used for audio data because they are easily read by computers and this

is often the data used for popular music. In certain regards, images and symbolic forms are similar; image data refers to images of scores, while symbolic data is a format that a computer can understand, such as a score notated in Finale or some other notation file. These different types of data have specific limits and uses. For example, metadata, which was explained above, is used in all search engines that look through bibliographic data. Audio on the other hand is not as easy to search but is very easy to obtain in standard MP3 format.

According to Burgoyne et al in Chapter 15 of the 2016, *A New Companion to Digital Humanities*, audio data is difficult for feature extraction—when a user aims to identify a particular query—because it comes in the form of large files. Historically “query-by-humming” (Burgoyne et al 2016) has been a popular MIR for feature extraction if it has been properly annotated. For query-by-humming, a user hums a tune in a microphone and the tune is matched with a piece. This, however, is by no means a complete picture of what audio can be used for. If an audio recording could be transferred to symbolic data, it would be more useful to MIR (Burgoyne et al 2016).

Symbolic Data, often is in the form MIDI or a readable score format, is easily recognizable by a computer and is used for information retrieval, classification, music performance, and music analysis. A symbolic form can retrieve sets of pitches (together making themes), rhythms, harmonic progressions, and more. Classification using symbolic formats identifies stylistic “emblems” such as a specific harmonic progression or the usage of specific intervals. This emblem is a defining characteristic. In terms of music performance, symbolic data is also used for expressive timing studies. Finally, for music analysis, symbolic format is used for automated analysis (this also overlaps with optimization) and for pitch spelling when MIDI is used (Burgoyne et al 2016).

Image data, like audio data, is difficult for a computer to recognize, and at present there is no consistent recognized form for sharing it. A score itself can be transcribed or turned into a MIDI format but that is time consuming. Optical Music Recognition (OMR) was created to solve this issue. OMR is a tool that can identify musical characters much like Optical Character Recognition can identify letters in typed images. This renders score images readable by computers (This will be further discussed in the section on SIMSSA, Single Interface for Music Score Searching and Analysis).

MIR is a multifaceted, multicultural, multidisciplinary tool. There are also seven facets of music information (Downie 2003):

1. Pitch
2. Temporal
3. Harmonic
4. Timbral
5. Editorial
6. Textual
7. Bibliographic

In the following graphic, I have given a representation of the overall shape of MIR, as it currently stands. The reader will note the breakdown into two large parts, locating and analytic/production, as discussed above. And within these, the reader will find the various of these fields as described above.

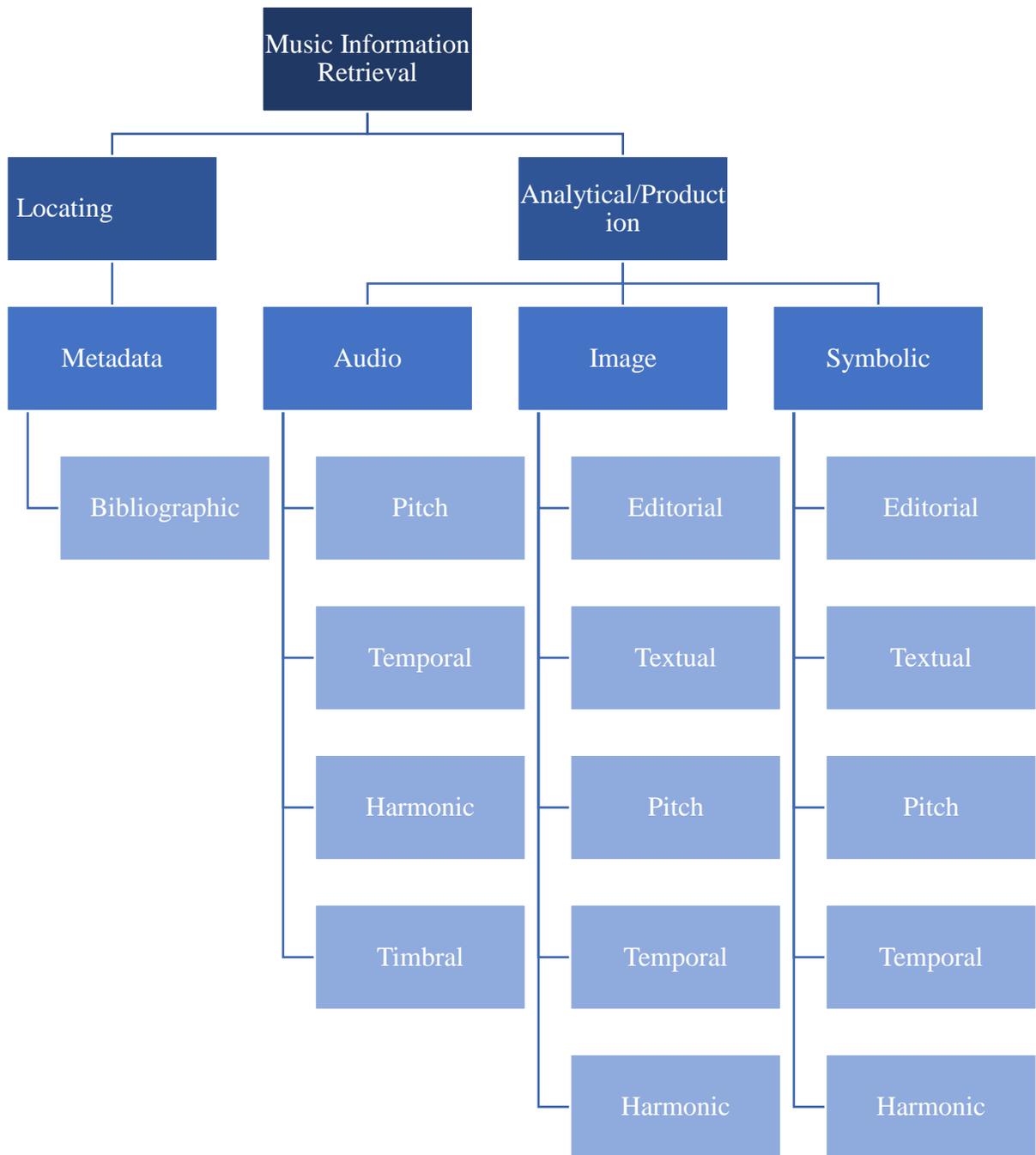


Figure 2 The second row and the last set of facets are the two categories of MIR system explained by J. Stephen Downie in his 2003 article. The four types of data are from chapter 15 by Burgoyne et al 2016

Though the graphic looks as if it represents a concrete situation, these lines are blurring due to changes since the turn of the century. These changes are being examined by ISMIR, the

International Society of Music Information Retrieval, and MIREX, the Music Information Retrieval Evaluation eXchange (Burgoyne et al 2016), but, as stated in their names, they only look at MIR tools (this is one of my five critical issues). This graphic representation has been included as a comparison point for the rest of Chapter Two, so I will be referring these types of data (Metadata, Audio, Image, Symbolic), facets (Pitch, Temporal, Harmonic, Timbral, Editorial, Textual, Bibliographic), and systems (Locating, Analytical/Production).

2.2 The MIR Tools

In this part of the thesis I shall look at several tools in MIR. Some of which are to be used by researchers in MIR and others for layperson use. First, I start with VocalSearch, which is now unavailable online but gives valuable information to the thesis. Next, I discuss three Single Interface for Music Score Searching and Analysis (SIMSSA) tools: Search the Liber Usualis, Cantus Ultimus, and Electronic Locator of Vertical Interval Successions (ELVIS). Finally, I examine a smaller tool which is Donnelly and Sheppard's Bayesian Network Algorithm that investigates timbre identification.

2.2.1 Vocalsearch

Vocalsearch is a web-based tool which was available to everyone and is used to identify unknown songs without metadata (Pardo et al 2008). Metadata is the information about the song such as title, artist, album, etc (Burgoyne et al 2016) and, without it, it is difficult to identify a song (Orio 2008). Vocalsearch was created by teams from University of Michigan and Carnegie Mellon University (Birmingham, Dannenberg, and Pardo 2006). I have chosen to include it as a

tool that is ‘complete’—as research grows this project may change, but it is a complete database when compared to the tools that follow in my discussion. This tool lets the user search—by humming a segment, by providing music notation, and by providing lyrics—using Melodic Music Indexing and Query-by-Humming technology.

Melodic Music Indexing is a way for the computer to understand the melodic content of a song. A song is annotated with the melodic content; often this is done through MIDI sequencing. MIDI is easily understood by a computer because it gives both pitch and duration. When a query is hummed, the computer matches it to the corresponding song. Song matching is problematic. Often, when a query-by-humming platform does not work, it is because the user did not hum the melody clearly or chose a different song layer, perhaps another instrument or vocal line (Dannenberg et al 2007). The tool must also equalize and understand the query, and, for Vocalsearch, this is done using a probability algorithm (Birmingham, Dannenberg Pardo 2006). The approach measures the similarity between the MIDI and the sung query for the large database.

Within MIR, Vocalsearch builds upon the existing audio data recognition and locating systems. It lets a specific song or number of songs be located using various queries recognized both through a typed search and a hummed audio search. Vocalsearch uses usual metadata searches if needed but seems to be more useful for unusual queries like, humming or notational search. The database itself is used for music with a lyrical content, hence the name, but the site is now unavailable, so the data from a user’s perspective is limited. A common issue with a database is that music is constantly being created, but this database of music will keep growing because a user can add songs (Pardo et al 2008).

2.2.2 SIMSSA

As I mentioned above, the in-progress tool is a research project called the Single Interface for Music Score Searching and Analysis (SIMSSA). In this section of the thesis, I describe three SIMSSA projects: “Search the Liber Usualis,” “Cantus Ultimus,” and “ELVIS.” These all have different goals and technologies, so including all three gives a well-rounded view of what goes into a tool.

2.2.2.1 Search the Liber Usualis

The Liber Usualis contains valuable information for those working on early music. The text is over 2000 pages, so it is difficult to locate the needed information. To solve this problem, SIMSSA decided to render its contents searchable and make it all available online. This tool lets researchers search the text for pitch sequences (either transposed or exact), neumes, contour, intervals, and, of course, text (Search the Liber Usualis Website is located at liber.simssa.ca). To do so, SIMSSA has used Optical Text Recognition (OTR), sometimes referred to Optical Character Recognition (OCR), and Optical Music Recognition (OMR).

OMR, as previously mentioned, is a computer method involved in “turning musical notation represented in a digital image in a computer-manipulable symbolic notation format (Vigliensoni et al 2011 423).” Using OMR with neumes, or square-note notation, is difficult because it is a precursor to standard musical notation. Because this notation is a precursor, there is no standard notation software, so the tool must translate the square-note notation to the standardized one. OMR must be configured to translate the first notation to the required notation. The translation to standard notation requires computer understanding of eleven neumes. SIMSSA

decided to use the ‘Music Staves Gamera Toolkit’ as a bank of algorithms to perform an analysis on 40 test pages of the *Liber Usualis*. The test pages were manually classified and annotated to double check the output of the algorithms. The algorithms used did the following tasks: created the staff lines, removed the staff, added ledger lines, and classified the types of neumes. When classifying neumes, the algorithm did not work 100%, so the final version was examined by a human to ensure perfection. These algorithms, however, do not tackle clef recognition and note identification.

Note Identification was made possible using horizontal projection of neumes, but this only worked for a subset of the eleven neumes. In conjunction with the algorithms used prior for determining types and placement of neumes relative to the staff, the starting pitch of the neume was identified using the average size of the neume and its “center of mass (Vigliensoni et al 2011, 426).” The clef was then identified and each neume was given a pitch relative to the clef. This was possible because the clef is always the first neume-like image in the line. The remaining set of neumes often have multiple pitches, so they were treated as exceptions to the above-mentioned method. These neumes were first split so the resulting output would correctly identify the multiple pitches. In conclusion, a different algorithm from the Music Staves Gamera Toolkit was used for each of the procedures, but, together, the algorithms rendered the scores from the entire book searchable.

The scores were made searchable using algorithms, then, the text was searchable through OTR technology in a simpler fashion to the scores. The “Search the Liber Usualis” project fits in the MIR chart above by being analytical and as a tool for locating scores and text. It is analytical because it uses an image of a text and looks at contour and interval, these being elements of

analysis and locating because it finds specific ideas based on the searched criteria. This is possible because of the computer's ability to 'read the music' once the algorithms translate it.

2.2.2.2 Cantus Ultimus

The "Search the Liber Usualis" can be seen as an initial test, laying the groundwork for the Cantus Ultimus. Their goals, however, are different. For the Liber, the goal was to make it searchable and make it easy for researchers to use the book. With the Cantus Ultimus, the aim is to preserve the ancient manuscripts digitally before they deteriorate further. The database shows images of the searched score, with typed lyrics, and standard notation on the side bar (Cantus Ultimus is located at cantus.simssa.ca/). Only a few sets of images have been added, but this project is still growing.

The Cantus Ultimus is part of SIMSSA primarily located at McGill University. This tool builds upon the existing Cantus Database with more digitized scores and Optical Music Recognition (OMR) technology. Researchers and plainchant enthusiasts can search through the database by text, genre, office, and by reference to the associated liturgical feast. Text queries include lyrics of the chant and the metadata for each. They can also make musical search using "Volpiano searches" which are searches using notes specifically. This can either be a normal search where A-B-C would show results for A-B-C, D-E-F, and any other series with the same intervals or a literal search where only A-B-C sequences would be shown (cantus.simssa.ca/).

Each query can yield multiple results, so, in effect, it is a locating system. The system locates based on notes, and lyrics, but, more importantly, it is an image searching database. The

ability to search through images was made possible through OMR and OCR with all of the algorithms used in the “Search the Liber Usualis.”

2.2.2.3 Electronic Locator of Vertical Interval Succession ELVIS

The Electronic Locator of Vertical Interval Succession (ELVIS) was created to give counterpoint the attention it deserves. In fact, a presentation on ELVIS, by Christopher Antilla, won first prize at the 2014 Montreal Digital Humanities Showcase and is funded by a Digging into Data Challenge award (located at <https://elvisproject.ca/>). The goal of ELVIS is to look at musical style in terms of changes in counterpoint (Antilla and Cumming 2014). ELVIS is a set of downloadable scores in a database, a web-based application, and a downloadable tool. These three aspects have taken many people to create it. Most of the people, such as Ichiro Fujinaga and Peter Schubert, are from McGill University in Montreal, those working on the harmonic side of counterpoint are headed by Ian Quinn from Yale University, and the University of Aberdeen has also been involved with this project. But, the software for the downloadable tool, music21, was created by Myke Cuthbert at the Massachusetts Institute of Technology (Music21)

Music21 is a python based “toolkit for computer-aided musicology” (music21) that allows the user to search through any imported scores using basic programming language. What this means is, by using commands such as *if x then y*, then a desired output can be found. This works especially well for big data queries in MIR (Antilla and Cumming 2014). Using the ELVIS database, the scores can be imported and searched using music21. The scores in the database can be searched through the ELVIS website and, using the web app, patterns are located. The Downloadable software is a VIS, Vertical Interval Succession—meaning a set of

harmonic intervals in a particular order—, framework used on music21 (ELVIS project). The framework uses *n-grams* when referring to the number (*n*) of vertical interval successions. This analysis uses intervals without quality instead of note names to compare many works regardless of key (Antilla and Cumming 2014). This software is used on Python, a standard programming language, so those with a knowledge of programming commands can get the most out of it. For those who do not have programming knowledge there is a Counterpoint web app (counterpoint.elvis.ca).

The application for ELVIS is called the Counterpoint Web App on their website (ELVIS project) and is specifically for pattern recognition. This web app continues to use a VIS framework, but it is more limited in query possibilities than the downloadable extension for music21. Getting to the application through the website is problematic because of a broken link or, perhaps, the web application is not finished. As previously mentioned, SIMSSA is building tools and many of the tools are still in progress.

Music Sonification is used in the ELVIS project to turn the music notation data into sound but can be manipulated by the researcher. Accessibility, in this case, was the main concern because not all researchers will have in depth knowledge of recording or sound mixing software. To solve this problem, the ELVIS team have created a graphic user interface. This is a graphic representation of music and the most useful audio tools for interval analysis. The concentration on interval analysis is because ELVIS is for contrapuntal analysis and pattern recognition (ELVIS project). ELVIS is both a locating and analysis tool. The locating part is from the web app because it only locates patterns. The analytical axis, however, is much more in depth and is available for a wide variety of early music using the VIS Framework and the programming language. Though the intention of the project was for counterpoint alone, the VIS Framework,

music21, and the use of *pandas* libraries—where the scores themselves are kept—make possibilities endless (ELVIS project).

2.2.3 Donnelly and Sheppard Bayesian Network Algorithm

Donnelly and Sheppard—researchers from University of Notre Dame and Montana State University respectively—found that timbre has not been fully explored in MIR, so they have modified an existing algorithm derived from Bayesian probability Networks. This new system of steps identifies different timbres in music. This can be used to establish another way of organizing and searching through music in a large corpus. In Donnelly and Sheppard’s article, “Classification of Musical Timbre Using Bayesian Networks,” nearest neighbour and vector machine as timbral identification models are compared to this new model. Upon comparison to the other models, the Bayesian algorithm better differentiates strings, but still has drawbacks. The other models better differentiate between aerophones, like woodwinds and brass, but, together, it appears the models can differentiate all instruments together. This seems to still be useful as a method for categorizing string instruments and, in conjunction with the other tools, can categorize all instruments.

The target audience for this method, are researchers and others who want to organize a database using instruments within a musical track. This can grow the locating section for audio as an alternative to metadata, but this would be for smaller tasks examining instruments. This is included as a smaller technology that has capabilities for MIR and to show the possibilities for connection between MIR and Optimization, which is the following chapter.

2.3 Critical Analysis

This chapter thus far has explained what each of the tools do. This section examines each tool critically. I discuss the assumptions made, and further extensions of the tool that were not examined in the articles themselves. I go through each of the tools in this order that they were previously presented, so first I examine VocalSearch, SIMSSA—Cantus Ultimus, Liber Usualis, and ELVIS—and finally the Bayesian Networks presented by Donnelley.

2.3.1 VocalSearch

VocalSearch takes audio input, which is difficult because audio input must be taken apart to match a specific line in a song. However, it is not mentioned if a melody sung in a different key from the original will match a song to the input. Though melodies are often remembered in the original key, the user may not have the range to do so. Also, this article does not mention the matching of a song from the database to a slightly inaccurate input, so it likely would not work in such a case.

VocalSearch achieves its goal of being able to reach a large audience using the internet and having multiple ways of searching queries. Setting up such a database takes a large body of songs, but to keep a database like this current, new songs must be added regularly. To do this, the makers of VocalSearch have included a function that allows users to add content to the database. There are a few issues with users adding content. As previously stated in the Introductory chapter, the errors made by a computer program are due to human error. This human error can be in the programming itself, but more often it is in the input for the program. As mentioned with VocalSearch, there are multiple methods of searching, so the person who inputs a song must

enter all correct information. If incorrect information is added, then the tool will not work correctly rendering it inutile.

2.3.2 SIMSSA

SIMSSA has multiple projects, so I will critically analyze each of the projects from SIMSSA. Overall, SIMSSA uses scores images and creates databases using OMR, OTR, and other technologies.

2.3.2.1 Search the Liber Usualis

Using optical text recognition (OTR) and optical music recognition (OMR) the *Liber Usualis* is searchable. Meaning that, by typing in a search bar, matching text or music is highlighted and, by using the colour coating available on the web-based tool, multiple searches can be highlighted at once. This is useful for researchers who need specific information from this 1000+ page text. More information on the tool can be found in section 2.3.

OMR and OTR are used when the file format has come from images and are, therefore, not searchable. These technologies make the document searchable by translating the image data into a format recognized by the computer. For OTR, this translates the image of a letter to the letter itself while OMR must attach the letter name and the function of note. This increases the margin of error. An issue I have found when using the tool is that coloured highlighting box around the searched content is not completely accurate. With some searched content, the box is around a set of words that do not contain the searched item. Also, an assumption made is that the user wants the entire sentence highlighted when searching for a specific word or group of words.

This calls into question how OTR works because if it turns a text searchable, then it should only highlight what is searched.

2.3.2.2 Cantus Ultimus

Cantus Ultimus uses digitized scores and OMR, to create an interactive and searchable score database. This not only gives a researcher the access to the database, but also lets them search the score in multiple ways. Furthermore, the database gives the researcher access to the manuscript image online with the typeface version in the righthand menu. For example, if there are neumes in very small writing on the score image then the right-hand menu will give the modern notation of the score.

Currently, there are only a few scores or manuscripts, so the obvious improvement is to have more scores. The process to add a score, however, is very long even using OTR and OMR because all scores should be checked. Because the manuscripts have aged, can be faded, or overall difficult for a computer to read, checking is imperative to a proper database entry. What could help are Machine Learning and Optimization models that are discussed in further chapters.

2.5.2.3 Electronic Locator of Vertical Interval Successions (ELVIS)

ELVIS gives counterpoint priority in research by combining a database with a web app and music21. The database gives the user access to a set of scores while the web app and music21 lets the researcher search through the scores. The web app is designed for a non-

programmer to find recurring patterns, and music21 has more features and the entire score can be searched using programming language.

This tool attempts to cater to both the programmer and the non-programmer by using music21, that is based on python—a common programming language--, and the web app. However, the web app only allows the user to find recurring features, so a non-programmer has limited usage with this tool. It is assumed that a non-programmer will only want to use this tool to find recurring features while they could, also, be looking for specific vertical interval successions, or a specific set of notes.

2.3.3 Bayesian Networks

As previously stated, this model gives timbre attention because it can be used to add in a search. This model is, however, limited in its ability to distinguish between aerophones, but can better differentiate between strings. To approach this problem, the tool must be combined with others to achieve greater accuracy.

The goal of this tool is to differentiate between instruments and, eventually, search through a database and render it searchable by instrument. Another way to approach this is to look at the metadata which often contains instrument data. Using an OTR-like algorithm, the metadata can be searched for contributing artists and musicians. This would render a set of works searchable by the contributors which will often contain the name of the instrument each contributor plays and, therefore, the set of works would be searchable by instrument.

This approach is useful specifically for works where the contributors' instrument is unknown, and the unknown instruments are stringed. Upon combining this method with other similar methods, the usefulness will increase because all instruments can be identified.

Chapter 3: Optimization

I use the term “optimization” to refer to the increase of output for less time and energy in music analysis—the optimization of effort so as to achieve a result. More specifically, this section will look at Preference Rule Systems (PRSs) and Probabilistic and Statistical Models. The goal in optimization is to understand and reproduce a human perception of an input. My goal is to show that, by integrating more mathematics and computer tools, analysis can be optimized. This term was inspired by its customary use in the areas of Calculus or Business, where the optimization of space and resources is described in term of optimization problems. In music, the term pertains to David Temperley’s progression in analytical approaches.

Temperley’s the *Cognition of Basic Musical Structures* (2001) took a preference rule approach to musical elements. For each element, a set of Preference rules were outlined for a computer tool to analyze a piece of music for information. Following this, Temperley took a few of the elements examined in the 2001 book and applied a probabilistic approach to them using Bayesian Probability—a term referring to extensions of the acceptance of Bayes’ Rule³—to match the approach of similar perceptual fields. The 2007 book, *Music and Probability*, aims to build upon the previous set of preference rules and move further in the research. This is the method of Optimization to be addressed here.

This section of the thesis will explain a previous way of approaching a problem and explain how a new method has helped to optimize the older one. Like both of Temperley’s approaches, there will be a section on organization by Preference Rules and a section examining Probability and Statistical models. In the Preference Rule section, Temperley’s approach will be

³ Bayes’ rule is expressed as follows: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ where probability is P and items A and B are distinct and different. Upon acceptance of this theorem, a branch of probability is built called Bayesian Probability

discussed first. Following this, other preference rule methods and computer tools will be presented as they relate to Temperley's *Cognition of Basic Musical Structures* (2001). The second section will show various approaches to music analysis that involve different aspects of Probability and Statistics. Some of these approaches, like Temperley, use Bayesian Probability, and others concentrate on statistical analysis. Though the two sections are split in this thesis they are related since the hierarchy built in a PRS carries through into Probability. I have separated them in the thesis to better explain how a newer model has built upon Temperley's work but they are related. This is represented graphically in figure 3 where the dashed line represents the implicit link between the two main sections, even though they are distinct in their principal focus (i.e. a PRS or application of probability). The items under each of the main headings are the topics that are covered in this chapter. Bayesian Probability can encompass all of the subheadings under Preference Rules, but Harmonic Vectors and the application of Bioinformatics later in the chapter relate more specifically to other subheadings. This is also represented through dashed lines.

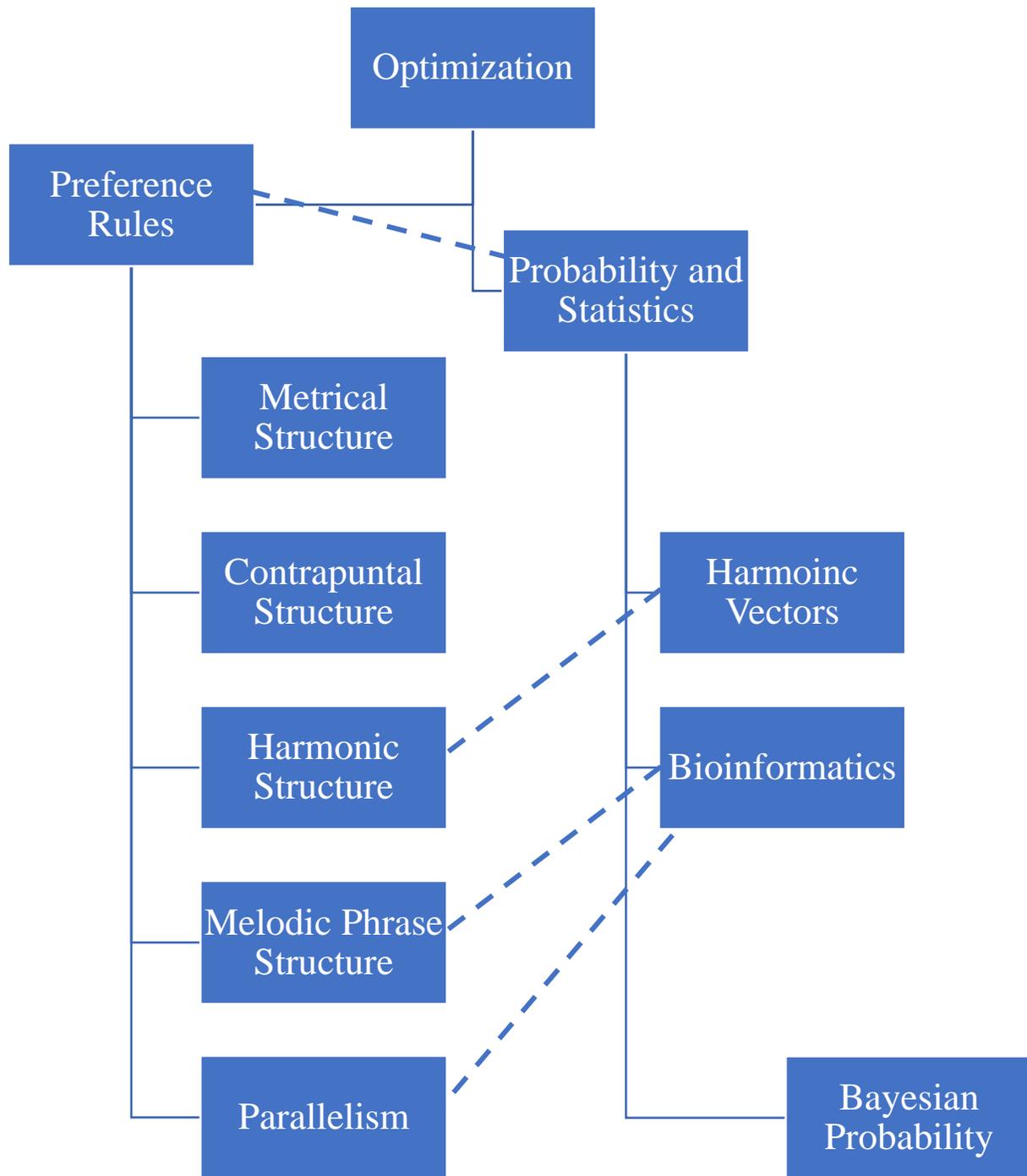


Figure 3 This is a graphic representation of the aspects of the field I concentrate on. It shows that Preference Rules and Probability and Statistics are not completely separate from each other.

3.1 Preference Rules

This section on Preference Rules will start by outlining David Temperley's Preference Rule Systems (PRSs) from *The Cognition of Basic Musical Structures* (2001). I concentrate on the first section of the book. Temperley uses a piano roll input for the computer and, based on the subsection in question, specific tests are performed to examine the usefulness of the approach. The subsections of this book—Metrical Structure, Melodic Phrase Structure, Contrapuntal Structure, Tonal-Pitch-Class Representation, Harmonic Structure, and Key Structure—will serve as subsections of the following chapter. Parallelism is the final subsection in this chapter and it was added because of a 2002 Temperley and Bartlette article, "Parallelism as a Factor in Metrical Analysis," that further explains the importance of parallelism (This article also gives a broader definition to parallelism which is important to further research). For each subsection, Temperley's findings from 2001 will be presented followed by the research that has built upon the findings.

In this part of the thesis, I take Temperley's model and examine how the next 16 years of research has built upon it. I will present a set of the comparable models and give a brief explanation of the element of Temperley the model builds upon. Following this section, I will critically examine the newer models and tools through comparison. I begin, however, with Temperley's 2001 book *The Cognition of Basic Musical Structures*.

3.1.1 Metrical Structure

As David Temperley explains in *The Cognition of Basic Musical Structures* (2001), the computer must concentrate on that beat induction when examining metrical structure. Beat induction is when the computer must understand or tap the beat. In some senses, the term refers to a ‘foot tapping’-like induction, but for the Temperley PRS it is for inferring meter. The meter is shown in a Lerdahl Jackendoff graphic model with different hierarchies of beats as shown in figure 4. This is a Metrical grid for 2/4 time where the lowest set of dots indicates the eighth note lever (the division of the beat level), the middle set of dots are the main beat (1, 2, etc.) and the highest level is the strong beat (the downbeat).

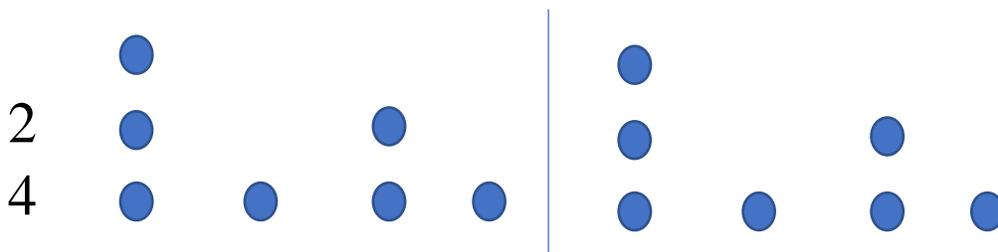


Figure 4 This is a beat hierarchy and described by Lerdahl and Jackendoff

For finding metrical structure, Temperley outlines the rules as followed:

1. Event rule: Prefer event on a strong beat onset
2. Length rule: prefer long events on strong beats
3. Regularity rule: prefer evenly spaced beats at each level
4. Grouping rule: “Prefer a strong beat at beginning of groups (Temperley 2001, 38)”
5. Duple bias rule: Prefer duple or triple levels (for example 3/4 instead of 6/8)
6. Harmony rule: strong beats align with harmonics change
7. Stress rule: prefer strong beats with loud events
8. Linguistic stress rule: prefer stressed syllables on strong beats

9. Parallelism rule: prefer the same metrical structure to the same segments

What these rules consolidate to is a set of preferences for a computer system to go through to find the “best-fit” for metrical structure. The computer will attempt to fit different meters onto a piece of music and chose a version where the most parameters are preferred. Because these are preference rules, in other words the computer does not have to have all of them true when choosing a meter, so the “best-fit” refers to the meter with most of the preferences.

Tempo is a bottom-up and a top-down process depending on how long someone listens to a piece of music in the same tempo. It is a bottom-up process because we need a few notes to perceive a tempo, but following these few notes it is a top-down process because we apply the tempo we have perceived to the music—as evident through foot-tapping, head bobbing etc. However, if the tempo were to change suddenly for expressivity, a person could catch it quickly. According to Desain and Honing, “beat induction is a fast process [since] only after a few notes a strong sense of beat can be induced” (Desain and Honing 1999,29), and, therefore, a computer inducing tempo is a large undertaking.

In Temperley’s writing he mentions the “most important work in the area of quantization (Temperley 2001, 27)” is a 1992 Desain and Honing study entitled “Time Functions Function Best as Functions of Multiple Times.” I mention this article because of its comparative approach and use of much the same rule-based models as Temperley. The 1992 article is a connectionist approach that uses stationary units and interactive units that change based on the surrounding material. The approach does not keep the length of notes the same but keeps the onset the same, which is problematic for Temperley. Even though this model offers multiple beats per time interval it cannot handle expressive timing (Temperley 2001)

The 1999 Desain and Honing study, “Computational Models of Beat Induction: The Rule Based Approach” used a rule-based model for beat induction of musical input and aims to explore the perception of tempo in people and in computers. The goal of this article is to look at rule-based models and provide an understanding of how these models create an initial beat structure. Desain and Honing examined the contribution and robustness of rules in different rule-based models. The important aspect taken from this article is that models, regardless of year created, can work more optimally with rules taken from other models. This points towards the mixing of rules and ideas which is in fact what Temperley has done to create his PRS.

Smith and Honing (2008) explains how the problem of expressive timing could be overcome. This study used rhythmically isolated segments—meaning that there was only rhythm as input—to incorporate expressive timing. This accounts for the fact that a person can easily change their original beat structure to incorporate expression. A technique based on Morlet Wavelengths was used to do so because of its similarity to human hearing and prediction⁴. This remains consistent with the overall goals of Optimization, which is to explain with greater and greater efficiency perception and human signal processing. These wavelets, however, are best used for short bursts of input similar to that of expressive timing at the ends of phrases.

The article first looks at the analytical techniques and the application of Morlet Wavelets to create a continuous wavelet (one that uses expressive timing). A wavelet is a representation of the repetitive rhythmic structure, such as a repeated rhythm or time signature. Then it puts the rhythmic findings into a hierarchy. Following this, the article finds the “foot tapping rate” (Smith and Honing 2008, 83) which is the basic tempo and, finally, the model is complete by showing

⁴ Definition taken from an Online Dictionary on time frequency. https://cnx.org/contents/Skft37_1@2/Time-Frequency-Dictionaries

the incorporation of expressive timing (Step 1 with Step 3). Overall, this model will provide an accurate analysis of foot-tapping. It will be further discussed in Section 3.2.

Hardesty in 2016 goes a different direction in building upon Temperley as well as Huron and Lerdahl and Jackendoff's *A Generative Theory of Tonal Music* (1983). His approach aims to identify rhythmical features and examine music prediction from the rhythmic and parallelism point of view. This will be further discussed in the parallelism section.

3.1.2 Contrapuntal Structure

As mentioned in Chapter 2 with the ELVIS project, counterpoint often does not get the attention it deserves. Temperley examines counterpoint with the goal of understanding the perception behind it. It is worth mentioning that the PRS for contrapuntal structure is geared towards a piano roll representation of a piece. Temperley uses the concept of "streams" which are a group of ideas in the same voice with minimal white squares. The white squares refer to moments of silence. Temperley's Preference rules are as follows:

1. Pitch Proximity Rule: prefer to avoid large leaps in a stream
2. New Stream Rule: prefer the least number on streams
3. White Square Rule: prefer the least number of white squares in a stream
4. Collision Rule: prefer cases where a square is in one stream
5. Top Voice Rule: prefer a single voice as the top voice, so there is minimal voice exchange

I would like to clarify that a stream does not refer to a phrase because, in contrapuntal structures, a stream can have multiple phrases. For example, one voice in a 4-part fugue would start with the

melody which can be multiple phrases, then the same voice will play contrapuntal variations with multiple phrases; this voice acts as one stream

A 2015 Komosinki article examined analysis of counterpoint for compositional research by using a method called “dominance relation.” This is a method that uses multiple criteria to do analysis like a PRS. It specifically looks at first species counterpoint and can produce an output of a composition. Because this is a composition tool, I will concentrate on the evaluative module of the method. The model will first always generate the first species counterpoint, but each item is evaluated by the following criteria:

1. Direct motion
2. A repeated note
3. A vertical imperfect consonance
4. A skip
5. A vertical perfect consonance reached by direct motion
6. Skips by tritone or larger than P5 except m6

These criteria are examined through the generated piece and they are all counted.

The output produced by a dominance relation will be either “dominated” or “non-dominated.” Using rules based upon the counterpoint method of Fux (Fux 1965), dominated counterpoint will have another counter point that is ‘better’ and this evaluation will repeat until a final, non-dominated counterpoint is found. This article builds upon Temperley’s rules but only in a general sense. Temperley’s rules are used to narrow down choices and find the best fit, while this method tests all rules on each counterpoint, and eventually finds the counterpoint that most exemplifies the rules.

Giraud et al in 2015 builds upon research on fugues. The input has the voices in the fugue already separated. This is much the same as Temperley's streams and uses "generic MIR techniques" (Giraud et al 2015, 79). I have decided to put this into the Optimization section for two reasons. First, it is an example of work lying between optimization and MIR. Secondly, it acts more as an Optimization tool than an MIR tool because of its small scale. The goal is not to create a database. Instead, the goal is to be used as an evaluative model for fugues.

This tool needs input that is already separated for computer use, so it uses files from the Humdrum toolkit because they have been previously separated into voices. This method concentrates on using tools to examine pattern repetitions and gives a complete analysis. It does so by identifying the subject, and countersubject(s), the key for individual occurrences, harmonic sequence, cadence, pedals, and overall structure. Giraud et al tested this method on 36 Bach and Shostakovich fugues. They found that, for some pieces, the analysis was complete and correct, but the method still gets false positives. Other results were completely unusable, but these were mostly double and triple fugues. More specifically, if the subject was correctly identified the overall analysis was more correct. Like any computer method, this one can be made better and Giraud et al makes suggestions on how. To make this optimal, Giraud et al suggests that the current method can be combined with probabilistic models. Probabilistic models will be discussed in the following section.

3.1.3 Tonal-Pitch Class Representation and Harmonic Structure

Tonal-Pitch Class Representation is important to the PRS of Harmonic Structure. The term Tonal-Pitch Class is taken from Temperley and I have understood it to mean the set of pitch classes creating a tonal structure (i.e. key area). Tonal-Pitch Class representation is the sorting of

the pitches in a piece to a specific key. The Preference rules outlined by Temperley are as follows:

1. Pitch Variance Rule: prefer to label such that nearby events within the same key
2. Voice-leading Rule: events a half step apart are preferred to be different letter names
3. Harmonic Feedback Rule: prefer a Tonal Pitch Class where the harmonic structure is good (meaning that there is a logical progression)

These rules help to decide a specific key and minimize notes outside of a chosen key. All keys would be tested for a specific idea and the best-fit would be chosen. The PRS for Harmonic Structure builds upon this assignment by adding roots and chords to the piece. These rules create a hierarchy of possibilities for the individual chords and, because the last rule for Tonal-Pitch representation considers harmonic progression, the progression is relatively accurate. This does not eliminate the analyst, however, because this is not 100% accurate. The PRS for Harmonic Structure are as follows:

1. Compatibility Rule: Prefer roots in the following order-> 1,5,3, flat3, flat7, flat5, flat9, ornamental (all others)
2. Strong Beat Rule: prefer chords on strong beats
3. Harmonic Variance Rule: prefer the next root to be on the circle of fifths
4. Ornamental Dissonance Rule: [ornamental dissonance is “if it does not have a chord-tone relationship to the chosen root] Prefer ornamental dissonances where the next or prior note is a tone or semitone away and/or on a weak beat

The PRS for Harmonic Structure still considers chords that are not part of the original key, and thus modal mixture and other temporary key changes are possible. This method also considers proximity, so modulation can be addressed.

To add to this, De Haas et al in 2013 created HarmTrace which stands for Harmonic Analysis and Retrieval of Music with Type-level Representation of Abstract Chord Entities. This tool is useful for tonal works to separate data using harmonic similarity estimation, chord recognition, and automatic harmonization. To explain further, this tool can recognize chords and show that different aspects of a piece are similar because of the harmonic structure or progression. This tool can do so by taking all the chord possibilities into consideration for the specific beat and extracting the most correct one. (The tool can also harmonize a progression which is useful for the performer, but not within the scope of this paper.) This article was included because it furthers Temperley's PRS: it can provide the automatic harmonization and similarity estimation. It does not need the previous Tonal-Pitch class representation PRS to figure out the specific chords. Instead it puts the possibilities into a hierarchical structure. The authors claim that this model can be used for MIR because it moves beyond theoretical uses and is practical as an internet-based method (De Haas et al 2013).

3.1.4 Melodic Phrase Structure

Melodic Phrase Structure is involved in multiple levels of a piece because melody itself often adheres to specific rules and works with other musical structures such as meter and harmony (Temperley 2001). Thus, Temperley's PRS must take all of these into account to be accurate.

The rules are as follows:

1. Gap Rule: prefer boundaries either at time between intervals or at a time at a rest before and interval
2. Phrase Length Rule: prefer 8 note long phrases

3. Metrical Parallelism Rule: prefer phrases that start at the same point in the metrical structure

The first rule refers to the time that could be between phrases or in a phrase. The Gap Rule is to make phrase boundaries at a rest or after a longer note value because these are both possibilities. An extension of this model will be discussed in 3.1.5 parallelism.

3.1.5 Parallelism

In *The Cognition of Basic Musical Structures* (2001), parallelism was mentioned and treated, and was revisited in Temperley and Bartlett's 2002 article. Parallelism was redefined as follows:

- a) Parallelism: repetition either exact sequence or contour
- b) Parallelism rule: “prefer beat intervals of a certain distance to the extent that repetition occurs at that distance in the vicinity” (Temperley Bartlett 2002, 134)

This twofold definition kept the existing definition but added contour and sequence in essence.

Emilios Cambouropoulos, from Aristotle University of Thessaloniki, in 2006 explored parallelism and melodic segmentation using a computer. Cambouropoulos wanted to incorporate parallelism into this method because it is often forgotten by analysts and has an impact on parsing data. Cambouropoulos used the pattern boundary strength profile (PAT) and the Local Boundary Detection Model (LBDM) to find phrase boundaries that take parallelism into account. PAT was first only able to extract patterns that are exactly the same, but Cambouropoulos modified it to extract patterns that are similar. The goal of this modification is to provide a more general application of parallelism which is exactly what Temperley wanted to do with the modification of his prior definition. Cambouropoulos was able to create a basic method for

melodic segmentation that incorporates parallelism, but it is not perfect as it does not provide the final segmentation of the piece.

As previously mentioned, Hardesty in 2016 published an article on music prediction and generation for rhythm. This method was based on finding parallelism, Lerdahl and Jackendoff's 1983 publication –*A Generative Theory of Tonal Music* (1983) –, and the psychological understanding of music. The psychological aspect of rhythm is based on “rhythmic anticipation and parallelism” (Hardesty 2016, 39). This method was only conducted on binary rhythm where strong and weak beats alternate, so the assumption is that an attack on a weak beat is followed by an attack of the strong beat. The method takes derivation of a rhythm to find the underlying operations to generate rhythms. The goal is to “[define] a collection of rhythmic building blocks (Hardesty 2016 abstract)” while taking psychological aspects of rhythm and meter and parallelism into account. The result is a hierarchy of rhythms based on duration. An interesting point is that the final outcome can still be the same if the input is different so long as they are derived from the same rhythm.

3.2 Probabilistic and Statistical models

Though this is a separate section from Preference Rules, Probability and Statistics encompasses the same hierarchical structure as a Preference Rule System. Often in Computer Music Analysis, different methods are layered to create an optimal outcome. The incorporation of Probability and Statistics stems from Temperley's move away from PRSs to a model that is more similar to other fields studying perception.

3.2.1 Introduction

In 2010, the Journal of Mathematics and Music published a special edition examining the first movement of Brahms' String Quartet in C Minor Op. 5, no. 1 to show different perspectives on Computer Music Analysis (referred to in the article as "computer-aided analysis"). The edition brought to light three major developments I explore further: Music Information Retrieval, Optimization, and Machine Learning. This section, however, will concentrate on Optimization in terms of probability and statistics. This will touch on work by David Temperley, Philippe Cathé, and Darrell Conklin. I will also introduce a method of using probability to assist in MIR, introduced in the previous chapter.

Temperley sought to improve Preference Rules with Bayesian Probability because it can do the job of preference rules. Preference rules are not used in other perception relation fields, like linguistics, so Temperley took their methods and adapted it for music. Temperley changed from a preference rule approach to a more generative approach using Bayesian Probability, which stems from the accepting of Bayes Rule as correct. This is when the probability of another event happening changes based on the occurrence of a previous event. Combining his previous work with that of *Music and Probability* (2007), Temperley created Melisma Version 2.0, available online for analysis.

Philippe Cathé located at L'Université Paris-Sorbonne looks primarily at Harmonic Vectors and uses a computer to perform the statistical analysis. The computer, however, does not perform the analysis itself, but, instead, treats each as a data file. Cathé attempts to keep the music in mind by, after the statistical analysis, explaining the interaction between the music and the vectors. With Harmonic Vectors, the changes can be heard in recordings making the statistical analysis seem more factual.

Darrell Conklin also employed probability, as well as bioinformatics for efficient pattern recognition. Finding patterns is an integral part of analysis but becomes subjective when choosing patterns for study. The goal of Conklin's work is to create an algorithm to find the *distinctive* patterns, which are patterns frequent within the piece, the corpus, and infrequent in a selected set of pieces, the anticorpus. This gives the analyst a set of patterns that may be important.

3.2.2 David Temperley's use of Bayesian Probability

In the *Cognition of Basic Musical Structures* (2001), David Temperley created a set of Preference Rules inspired *A Generative Theory of Tonal Music* (1983) by Lerdahl and Jackendoff. Similarly, *Music and Probability* (2007) takes a generative approach and combines it with Bayesian Probability. The reason for using probability was to use similar tools to language and vision because preference rules were not being used in these similar domains. Bayesian probability is a subset of probabilistic rules where the probability of an occurrence is affected by the occurrence of a previous event. This subsection will concentrate on select chapters from *Music and Probability* (2007).

The approach to analysis here is to first do a probabilistic analysis of the Essen Folksong Collection to find the probability of various musical building blocks, such as meter, keys—both in monophonic and polyphonic music—, and melodic ideas. This analysis sets the parameters for the computer program, so that the rest of the pieces analyzed will have a higher accuracy. Using the Essen Folksong Collection⁵, the parameters are set, and the analysis is completed through a

⁵ The Essen Folksong Collection is a collection of 10,000 folksongs collected by Helmut Schaffrath. These are located at <http://essen.themefinder.org/>

generative process. A generative process works by finding a structure based on the surface content of a work and then generating a surface in multiple choices (keys, meter, etc.). After generating a surface, the program will decide which is the highest probability based on the underlying structure. This simplified method will now be explained for meter, key—both monophonic and polyphonic—, and melodic ideas.

Meter has been well studied prior to Temperley's work in *Music and Probability* (2007), so this model aims to build upon previous models with a generative approach. A 'metrical grid' is generated from the piece based on the parameters set from the remainder of the Essen Folksong Collection, but there are many different possibilities of metrical grids for any given piece. As noted above metrical grid refers to the graphic representation of beats, strong beats, and main beat divisions in three levels as shown in figure 3 (Section 3.1.1)

The following steps are used in creating the optimal grid:

1. *Decide time signature*: choices between duple and triple meter and the individual time signatures within each category
2. *Generate the tactus*: this is the middle or second level of beats and is based on the notes that are present (simultaneously with 3)
3. *Addition of upper level beats*: indicates the actual beat division and is the highest level in the metrical grid (simultaneously with 2)
4. *Addition of lower level beats*: indication of the subdivision required for the excerpt. This is the lowest set of points on the metrical grid
5. *Generate note onset*: solid vertical lines that indicate where the actual notes line up on the metrical grid. (not in figure)

After generating many metrical grids, the tool would test the probability of the onset, with the assumption that the grid was correct. It would then multiply the grid by the probability of the grid itself. This would yield a probability value of statistically less than one and the highest scoring grid would be selected.

Upon testing this model on multiple pieces, Temperley compared it to the software that previously used preference rules to find the best fit. The tests showed that the PRS was more accurate when compared to the Bayesian model. Temperley hypothesized reasons for this. The reasons for higher accuracy with the first model is because the perception of rhythm is based on harmony, note lengths, and parallelism as well. Longer note lengths most often occur on strong beats such as the beginning of the measure and the Bayesian Model at the time could not take that into consideration.

In creating a computer model that perceives key, the musical facets the mind isolates must be taken into consideration. A key, at least in monophonic pieces, is composed of both pitch proximity and range and Temperley poses the question “What kind of pitch sequence makes a likely melody?” (Temperley 2007). This, once again, is a generative process where all keys are tested, but there is no obvious starting point when examining key, so Temperley relies upon previous research on key-profile. The key-profiles are heavily based on the Krumhansl and Kessler 1982 experiment. The experiment asked participants to rate the degree to which audible pitches belonged to an established key and, from this, a correlation was created. This experiment was successful in major keys, but minor keys were problematic because there are multiple versions of a minor key. Temperley made the needed changes to the established key profiles to incorporate minor keys and began constructing a model using Bayesian Probability.

To construct a generative process for key finding, Temperley used the key profiles as a starting point. He did an overall analysis of the Essen Folksong Collection to find a normal distribution, or bell-curve, of the pitches. Following this, a pitch is chosen at random from the peak area of the bell-curve to construct a range profile around it, and then, it is combined with a proximity profile. All keys are tested in this way and the key with the highest probability will be chosen as the key for the melody. This same method for key-finding is problematic for polyphony. This approach takes the structure from the surface material, but the surface of a polyphonic piece is dense and contains notes acting as passing or neighbouring tones. When examining a piece, many notes are not the tonic of a scale, so this would skew most computer programs. Temperley aimed to overcome this obstacle by segmenting the piece on the assumption that pieces stay in the same key for a little while. This assumption is based on the perception-based concept of ‘inertia’ where there is a lack of movement in an item (Larson 2004). In this case, it means that the key will stay the same for the amount of time affected by inertia. This also helps with the second problem of modulation.

Modulation occurs when a new key is introduced for an indeterminate amount of time. This is difficult for computer because two, or more, notes act as the tonic at different times in the piece. In the case of polyphony, this is overcome by segmenting the piece into smaller sections, as is already needed to look at polyphonic works. The smaller segments will show a higher probability to one key and a section that modulates will show a higher probability of another key. The segmentation, in turn, will assist in both, identifying modulation and key-finding in polyphonic works.

Melodic ideas in this case often involve expectation or error detection where the model attempts to answer this question: ‘does this pitch work in this sequence?’ Pitch expectation is

tested in two ways. The first is if the participant expects a pitch and the second is whether a participant can add a pitch.⁶ Temperley is concentrating on the first type of test and uses the Cuddy and Lunney (1995) experiment where participants rated the ‘fit’ of the next note in a corpus, not the Essen Corpus, from one to seven. The numbers were converted, by Temperley, into values to use the probability model. The values were used to test the strength of the fit of the note to explore the capabilities of the computer tool and to examine pitch sequence. Here, Temperley realized that the parameters work best if they were created by other pieces from the same corpus. The strength of best fit is much higher (from 0.729 to 0.870 in terms of correlation coefficient), but this shows that the computer tool does not work equally for all music but can give some insight.

3.2.3 Statistics and Harmonic Vectors

Harmonic Vectors is a newer harmony theory influenced by Riemann that aims to take a generative and systematic look at tonality that can be used for statistical analysis (Meeùs 2003). Nicolas Meeùs used this term from 1988 and wrote extensively on it into the twentieth century. My primary source for background information on Harmonic Vectors is a 2003 Meeùs article entitled “Vecteurs harmoniques.” This takes the motion of scale degrees and systematically sorts them into either Dominant (V) or Subdominant (SD) Vectors. The two types of vector are based on classification of progressions from Schoenberg and Sadaï, who wrote an extension of Schoenberg’s work. The reason for this analysis is the assumption that a chord alone has no meaning but creates its function within a succession of chords; therefore, the meaning is

⁶ Temperley refers to this as either the perception paradigm or the production paradigm

generative. These vectors can be graphically represented and can be used for statistical analysis but may not be representative if done on few works (Meeùs 2003).

Philippe Cathé took Harmonic Vectors and combined it with Computer Music Analysis to dig deeper into a set of works. There are three levels of research with Harmonic Vectors: finding regularities, finding pendulums, and finding correlations between the other two levels and the music (Cathé 2010a). Cathé expands on vectorial pairs (Meeùs 2003), an analysis looking at the pairs of side-by-side vectors, and mono vectorial sequences, meaning the same vector repeated, as methods for finding regularities. Pendulums help to further differentiate composers based on their vector use. A pendulum is a series of three vectors where the first and third vector are the same and the second vector is different. The final level of research brings back the music and aims to find correlations between the music and vectors found. The goal is to understand *why* a vector is used (Cathé 2010a). These three stages help to further explore a set of works.

The application of harmonic vectors for statistical analysis was mentioned and used by both Meeùs and Cathé. Both expressed the analysis in a table of percentages, organized by movement of scale degrees, the types of vector, and level, or with graphic representation, as line diagrams or graphs. The diagrams express the amount of each vector (Meeùs 2003), vector pair, mono-vectorial sequences, or pendulums (Cathé 2010a), most often in percentage, and break this down by era and composer. The computer has assisted Cathé in the three-level analysis by cutting down on the time and making the output as unbiased as possible. To perform comparisons, Cathé uses ‘Charles.’ ‘Charles’ is a computer program based on Microsoft Excel that gives proportions vectors (pair, pendulums, etc.) for a certain piece or a set of pieces, or data files. The output is expressed most often in charts or linear graphics. This gives the analyst

another method to represent the data and makes comparison easier between eras, composers, and compositions.

The idea that works of music taken from different eras sound different is not new. Harmonic Vectors aims to show this through the change in proportions between eras. Each era has a different average of each vector, vector pairs, pendulums etc. that can be identified through larger scale comparative analysis (Cathé 2010b) and represented in the form of statistics. In addition to eras, a comparative statistical analysis of harmonic vectors can also be applied to composers and compositions. All composers and compositions are slightly different, so Philippe Cathé took ten versions of *Vater unser im Himmelreich* and compared the usage of Harmonic Vectors (Cathé 2010b). A composer uses different amount of each vector (pair, pendulums, etc.) by piece, but the percentage remains very close (Cathé 2010a). This can also be used to show the degree of difference between two composers meaning that a composer's use of vectors is consistent by composer.

3.2.4 Distinctive Patterns using Bioinformatics and Probability

Looking for patterns is needed in all analyses and finding patterns that are distinctive is paramount. According to Darrell Conklin, a distinctive pattern is one which is frequent within the corpus when compared to the frequency within the anticorpus. The algorithm that was created aims to find the distinctive pattern within the corpus to narrow down the possibilities for the analyst (Conklin 2008). The corpus is a specific piece or set of pieces that are examined, so the distinctive patterns found is over-represented in the corpus. The anticorpus, on the other hand, is a piece or a set of pieces, often by the same composer, where the distinctive pattern is

under-represented. The frequency needed for distinctiveness, the corpus, and the anticorpus are all determined by the analyst. I will now explain a few applications of distinctiveness.

In this section, I will look at two different applications of this method done by Darrell Conklin. The first is on the Essen Folksong Collection and the second is on Johannes Brahms' String Quartet opus 51 no.1. The reason for choosing Conklin's application is to look at the approach of a researcher who commonly examines Music and Machine Learning (from Basque University Webstire <http://www.ehu.eus/cs-ikerbasque/conklin/>) and to further explain distinctiveness with an analysis. Both of the analyses use the similar following formula:

$$\Delta P \stackrel{\text{def}}{=} \frac{p(P/\oplus)}{p(P/\ominus)} = \frac{c^{\oplus}(P)}{p(P/\ominus) \times n^{\oplus}}$$

The middle expression (between the two equal signs) refers to the probability of a pattern in the corpus (\oplus) or in the anticorpus (\ominus). The last expression is used to find the value of ΔP , also known as likelihood of P ($I(P)$). The numerator is the total number of a pattern in the corpus and the denominator is the probability of a pattern in the anticorpus multiplied by the total number of events in the corpus.

The first analysis was conducted on the Essen Folksong Collection, the same collection used by Temperley in his *Music and Probability* (2007), and, more specifically, the Shanxi, Austrian, and Swiss folksongs. Conklin was searching for the "maximally general distinctive patterns," (Conklin 2008,1) which are patterns that can be used for classification but are not so general that they occur in almost all pieces. For a pattern to be considered interesting, or frequent, it must be in a minimum of 20% of the corpus. The likelihood ($I(P)$), also known as ΔP in later works, must be greater or equal to 3. This study showed that, for each region, there is a maximally general distinctive pattern that can be used for classification purposes (Conklin 2008).

The second analysis was on the first movement of the Brahms String Quartet, opus 51 no 1, and the anticorpus used was the string quartets no 2 and no 3. For the best comparison, Conklin only uses the first movement of no 2 and no 3. The goal was to show that the motives Forte found in his article “Motivic design and structural levels in the first movement of Brahms’s string quartet in C minor” (1983) are found as distinctive using this analysis, excluding two motives that cannot be maximally general. This is converse to when David Huron revisited the same analysis in 2001, where Huron found that only the alpha motive was distinctive (Conklin 2010).

I will now outline what was determined by the analysis. The minimum frequency, in this study, for a pattern is 10 and the likelihood of a pattern, renamed to the ΔP , is minimum 3 to be considered distinctive. The Humdrum kern formats were used for an easily available and computer compatible format. When the analysis was completed, all of Forte’s motives, not including the mentioned exception, were labeled as distinctive (Conklin 2010). This shows that the tool can be used to identify likely distinctive motives, but the analyst will still need to analyse the data for a complete picture.

3.3 Critical Analysis: Optimization

The chapter thus far shows the progression made in research in general and specifically that David Temperley made from *The Cognition of Basic Musical Structures* (2001) to *Music and Probability* (2007) by exploring the previous research, reasons for looking at probability, and the use of Bayesian Networks. In essence, the recent research in Optimization builds upon what Temperley provides or upon developments mirrored by Temperley. (Temperley has more recent publications, but these will be discussed in the conclusion of the thesis.)

3.3.1 Preference rules: Metrical Structure

The Smith and Honing use of Morlet wavelets was discussed in 3.1.1 as a method to incorporate expressive timing into beat induction. This method has its limitations. Firstly, the method does not work by exposing the tool to the music because the input must be in an isolated rhythm form. This means the tool cannot perform beat induction on a non-separated piece. Another issue is the selection of tempo is not as sensitive as needed. This method has made leaps and bounds in testing and creation but cannot currently work as a stand-alone program. And, because of its current limitations, the method cannot be a simple online application at this point, so it is only useful to a small number of people.

The first improvement is to make it either a stand-alone program or an addition to another larger tool. As its own stand-alone program, it would have much use to a researcher but may be a teaching aid for a student to learn expressive timing or beat induction. A more wide-spread use of this tool would be in playback software for scores to determine the efficacy of a playback. If the tool could not find the tempo of a piece as played in a playback, then it would show that the playback is not as similar to human playing. However, this tool does help to further the goal of Optimization by getting closer to human beat induction. In time, if work on beat induction continues, researchers may understand how people can find the beat and adapt it quickly.

3.3.2 Preference Rules: Counterpoint

The extensions of counterpoint from 2000 to 2016 have concentrated on the evaluative or compositional side, but they are still useful to analysts. The Komosinki article concentrates

heavily on composition but it gives an evaluative approach for the generated composition. On a smaller scale this tool is useful for evaluating a first species counterpoint which is taking an opposite direction than Temperley. It has been included to show a different use for Temperley's Preference Rules. It is useful to an analyst by giving a general outline of evaluative criteria needed by a computer. On its own, it needs to stay with a generative model because of the dominated vs. non-dominated output, but it is a good model for future evaluations of generative models.

The tool proposed by Giraud et al gives the analyst a strong head start on fugue analysis if the subject is properly identified. This tool is best used on a larger corpus of similar fugues (i.e. by the same composer in the same era) if it were to be combined with probabilistic models. The best probability of subject length, key notes used etc., is found when the corpus is evaluated independently. This was a trend in probability, because probability of certain gestures change based on the composer. This tool would indeed be best used in conjunction with a probabilistic model, but extra work needs to be done to separate a set of fugues into streams or voices. To separate the voices Temperley's preference rules to examine streams can be used if they are indeed one in the same. However, neither of these tools examine fugues with multiple instruments. This is left for further work.

3.3.3 Preference Rules: Tonal-Class representation and Harmony

HarmTrace can estimate the harmonic similarity, recognize chords, and automatically harmonize an input. This tool does not need a set of Tonal-Pitch Class rules or key profiles. Instead it uses a hierarchical structure to narrow down its choices. The authors of the article further say that this model can be used for MIR because it is practical as an internet-based

method. An issue that is not addressed is what kind of input can be used with HarmTrace this is one of my five Critical issues. If the input needs to be separated in some way then old Humdrum files could be used, but if there is an image score input then any clear scan of a score can be used. Another common input is a music notation software input (such as a Finale file), but these formats are specific to the notation software that is being used. Furthermore, an audio file input would be optimal because they are widely available, but this is not practical because no recording is perfect.

3.3.4 Melodic Phrase Structure and Parallelism

The PAT—pattern boundary strength profile—and LBDM—Local Boundary Detection Model—have improved with Cambouropoulos’s modifications in 2006, but since then parallelism has not been in the forefront of research. This more generalized application of parallelism is imperative for pieces where a repetition is ornamented or changed slightly, but it is often not considered with analysis tools because they often examine recurring features or one specific task.

Boundary detection is generally used for parsing data and by incorporating parallelism the boundaries are more accurate. By putting HarmTrace and PAT/LBDM together, the output could have a higher accuracy and can provide a precise parsing of data as needed for analysis. The final segmentation could be obtained for the PAT/LBDM outputs by using the HarmTrace harmonic infrastructure. This would be a way to leverage the strengths of both models to provide the user with a more complete outcome.

The Hardesty 2016 tool for examining rhythm has a strong basis in rhythm and music generation. It has further uses in optimization because it incorporates psychological elements,

however, the goal is not completely realized. The tool can only process and generate binary rhythm, but, with further research, the tool can come close to the human music prediction. Thus, it furthers optimization's goal in understanding how humans perceive rhythm and can predict it.

3.3.5 Probability and Statistics

The tools presented in the section on Probabilistic and Statistical models take three different approaches to using probability and statistics in Computer Music Analysis. Temperley looked at Bayesian probability, the set of probabilistic principals following the acceptance of Bayes' Rule, to incorporate his previous research in PRSs with cognition in similar fields to music. Cathé's approach aims to always keep the music in mind, so the computer looks at every data file, music in this case, and the analyst makes the final comparisons and assumptions looking at both music and harmonic vectors. Darrell Conklin takes bioinformatics and probability for finding *distinctive* patterns, and the method parses music giving the analyst the patterns that may be important.

Temperley's use of Bayesian probability is to be used in his online database. Overall, the generated coefficients can be used in other probabilistic models and in other corpus studies. As was stated by Temperley, the coefficients are more accurate when generated for a specific corpus, so for maximal accuracy this should be done. Furthermore, these coefficients can be used in any generative theory if they are based on the same corpus. This is also its limitation since re-analysing a set of works when investigating a different corpus is time consuming. This can sometimes defeat the purpose of a computer model as it does not save time and energy.

Overall these models take a set of data and provide an output of specific generalizations. For example, Cathé has generalized the percentage of use for each harmonic vector by composer,

meaning that each composer has a distinct percentage. This can be further combined with a study on authorship in 1963. This study was on literary works and measured the specific ratio of simple words such as upon, such etc. The amount some words were used is distinctive to the authors. The Poisson Process, a specific aspect of probability, was adapted to complete this method. This could potentially be adapted to music where, instead of words, harmonic vectors are used. This application is further discussed in the concluding section of the thesis.

Chapter 4- Machine Learning

4.1 Introduction to Machine Learning

Machine Learning can be defined as the process of teaching a computer (the machine) to devise new tasks, and in the case of music, to perform these new tasks on musical works. This has applications for many aspects of Computer Music Analysis, but the focal point of Machine Learning is the tool itself. The tool or method must provide a relatively accurate output on a first stage analysis so that, in turn, the tool can reliably produce correct output for other pieces. This differs from MIR and Optimization, because for MIR the goal is a database, and for Optimization, as I have described it, the goal is to understand and reproduce a human perception of an input.

Music poses many challenges for any computer-based analytical tool, and, as such, the analysis of full works of music using complex ideas is not common in Machine Learning. Machine Learning is used in multiple disciplines. When used for music, the input is often oversimplified (Widmer 2006). The field of Machine Learning as applied to music is still in its infancy. Thus, I can only give a cursory overview of some of its developments. (Recently, a special issue of the *Journal of Mathematics and Music* concentrated on Music generation in Machine Learning, but this is an exceptional development.)

In this section, I show several emerging possibilities for Machine Learning as well as precedents. I do so in an introductory manner because the actual processes of Machine Learning and their application are too complex to be treated exhaustively in a thesis of this scope. (I will discuss the literature of Machine Learning primarily from the angle of a music theorist although it holds considerable possibilities for other domains such as composition.) Unlike previous

chapters, the critical analysis for this chapter is in the conclusion of the thesis, since Machine Learning has importance to Computer Music Analysis as a whole.

4.2 Outline of Selected Tools

In this section, I aim to expose different tools in Machine Learning. First, I start with a tool that assists guitarists with ornamentation. The next two sections build upon one another as they are both created by Darrell Conklin and the second builds upon the first in terms of segmentation. It is also an application of the multiple-viewpoint system discussed in the Literature Review. The final tool is an analysis of analysis using Machine Learning. I concentrate on Kirilin and Yust's smaller details because it is one of the few Machine Learning models that directly adds to music analysis.

4.2.1 Ornamentation in Jazz Guitar

I begin with a recent development in the application of Machine Learning to music. For jazz guitar works, ornamentation is important because it is how expression is conveyed, but it is not written in the score. The performer must come up with the ornamentation themselves or go through countless recordings. Giraldo and Ramírez have attempted to address this problem with Machine Learning. This tool aims to take an “un-expressive score” (Giraldo and Ramírez 2016, 107) and add expressive ornaments to it. This machine learning tool uses 27 sets of audio input from a professional guitarist as a test set. Using a group of ornamentation vectors, the audio input was aligned with the score to create an expressive score of the recording. In effect, a non-expressive score was put together with a set of vectors derived from expressive scores. While the

primary goal of the study was to create a Machine Learning tool, a secondary goal of this tool was to give new guitarists an expressive score to read to help them learn the ornamentation practices.

Following the use of the test set, the tool was further tested on un-expressive input to get an expressive output. The output of the tool was a generated MIDI or other audio format recording that combined the un-expressive score with the Machine Learning ornamentation. The researchers determined that the overall stylistic and grammatical correctness of the tool is a statistical 78%. This tool does need further work, especially in refining itself as a Machine Learning tool. In terms of its secondary goal however, it does fill a void in jazz guitar performance.

4.2.2 Melodic Analysis with segment classes

Darrell Conklin's name appears frequently in machine learning as applied to music. His research centres around the problem of music as a multi-faceted entity. The article, entitled "Melodic Analysis with Segment classes" (Conklin 2006), is a stepping stone towards his later research that I will discuss in 4.2.3 (The basis for this article includes the Conklin and Whitten 1995 article discussed in the Literature Review). Conklin's 2006 article depends upon a concept called "viewpoints." The idea behind viewpoints is to take a cross section of musical structures and estimate the accuracy of the output. The aim of this study is to "demonstrate how the viewpoints representation for segment classes can be applied to interesting music data mining tasks" (Conklin 2006, 350).

Conklin's method is based on a study of natural language and its segmentation. For data mining, music must first be in a format understood by the computer and it must be hierarchal.

Accordingly, Conklin uses specific hierarchical and searchable terminology. A musical object is a note, a segment is a set of musical objects, and a sequence is a series of many segments in a specific order. Melody is a type of sequence: it is a set of notes in a specific order where the order of those sets is specified. Segmentation is a fundamental aspect of Conklin's analysis. There were two methods of segmentation tested. The first was phrase boundaries and the second was meter. Each test involved segmentation created using a viewpoint based on a set of pitches. The particular expression determined by Conklin is as follows: $\text{set}(\text{mod}12(\text{intref}(\text{pitch}, \text{key})))$. The method succeeded most with phrase and metric segmentation undertaken by beat (98%), note (92%), and bar (91%). (There was also successful interval level (94%) which was not segmented.) As is obvious by the percentages, the most successful was for segmentation by beat.

While the immediate task in Machine Learning is to create a tool, Conklin's secondary task was to discriminate style. The segmental viewpoint by beats can be used in future models for the secondary task. Conklin discusses the further work that needs to be done in this regard. Firstly, the length of segments must be examined for a corpus, meaning a collection of a *style* of music. Secondly, the problems of the automated segmentation, meaning the segmentation done by the computer, should be compared to human segmentation.

4.2.3 Chord sequence generation with semiotic patterns

Conklin's 2016 article, "Chord Sequence Generation with Semiotic Pattern," addresses the semiotic value in trance music—a type of fast electronic music, like techno, centred predominantly in Europe—when the latter is generated by a Machine Learning model. Aspects of the chords in trance music have intrinsic meaning and, therefore, the meaning must be kept to

have an accurate stylistic representation of the music. Conklin's model aspires to generate a chord sequence for trance music that keeps the qualities of trance music intact.

The semiotic patterns of trance music are defined as a sequence of "paradigmatic units" (Conklin 2016, 94). According to Conklin, the paradigmatic unit is when an idea is given a variable (a letter name) so that a pattern of these variables can be discussed. Viewpoints, a statistical model discussed above, is used to map, or create, an output according to a plan. Conklin's viewpoints are based on the following criteria: chord, root, chordal quality, inter-onset-interval (meaning the start and stop points of a particular sound), duration, chord diatonic root movement, chord quality movement, a combination of root and quality. Conklin describes the combination of root and quality as "crm. (cross product) cqm." The cross product is a common vector operation. This combination was chosen to generate the chord, taking into account the intrinsic meaning for trance music. I should note that Conklin only used a sampling of trance songs, so the results need to be further examined in terms of a larger trance corpus.

The goal in Machine Learning is the tool itself. Conklin states that the best algorithms, like the ones presented and other viewpoints, can be determined for a corpus. To further explain this, important aspects of a corpus can be identified, and the best algorithms can be defined and used like the "crm (cross product) cqm" used for generation in this article. Conklin also mentions that this method can be used for analysis.

4.2.4 Analysis of analysis

Kirlin and Yust's 2016 article "Analysis of Analysis: Using Machine Learning to Evaluate the Importance of Music Parameters for Schenkerian Analysis" aims to get a machine to develop the music theory branch of Machine Learning. The goal of the article was to create an

analysis of a score using a model resembling Schenkerian Analysis. While this goal was not realized, the article is still noteworthy because of what the researchers explored and the Machine Learning tool they created. Schenkerian Analysis involves reducing the work in question by finding patterns of ornamentation and elaboration. This task is difficult to teach a computer without stipulating the exact features to find. Kirilin and Yust defined eighteen features and then sorted them into categories. These became stepping stones towards creating a Machine Learning tool.

First a hierarchy of notes was created using a tool called a “maximal outerplanar graph.” Then the eighteen features were defined as they relate to the Left note, Middle note, and Right Note.⁷ The middle note has the following six features:

- SD-M The scale degree of the note (represented as an integer from 1 through 7, qualified as raised or lowered for altered scale degrees).
- RN-M The harmony present in the music at the time of onset of the center note (represented as a Roman numeral from I through VII or “cadential six-four”). For applied chords (tonicizations), labels correspond to the key of the tonicization.
- HC-M The category of harmony present in the music at the time of the center note represented as a selection from the set tonic (any I chord), dominant (any V or VII chord), predominant (II, II6, or IV), applied dominant, or VI chord. (The dataset did not have any III chords.)
- CT-M Whether the note is a chord tone in the harmony present at the time (represented as a selection from the set “basic chord member” (root, third, or fifth), “seventh of the chord,” or “not in the chord”).

⁷ These lists from pages 135-136 are shortened versions of the lists presented in Kirilin and Yust 2016

- Met-LMR The metrical strength of the middle note's position as compared to the metrical strength of note L, and to the metrical strength of note R (represented as a selection from the set "weaker," "same," or "stronger").
- Int-LMR The melodic intervals from L to M and from M to R, generic (scale-step values) and octave generalized (ranging from a unison to a seventh).

(Kirlin and Yust 2016, 135)

The left and right notes together have the following twelve:

- SD-LR: scale degree (1–7) of the notes L and R.
- Int-LR: melodic interval from L to R, with octaves removed.
- IntI-LR: melodic interval from L to R, with octaves removed and intervals larger than a fourth inverted.
- IntD-LR: direction of the melodic interval from L to R
- RN-LR: harmony present, as a roman numeral, in the music at the time of L or R
- HC-LR: category of harmony present in the music at the time of L or R, represented as a selection from the set tonic, dominant, predominant, applied dominant, or VI chord.
- CT-LR Status of L or R as a chord tone in the harmony present at the time
- MetN-LR A number indicating the beat strength of the metrical position of L or R. The downbeat of a measure is 0. For duple or quadruple meters, the halfway point of the measure is 1; for triple meters, beats two and three are 1. This pattern continues with strength levels of 2, 3, and so on.
- MetO-LR A number indicating the beat strength of the metrical position of L or R as an ordinal variable and treated differently in the algorithm
- Lev1-LR Whether L, M, and R are consecutive notes in the music

- Lev2-LR Whether L and R are in the same measure in the music
- Lev3-LR Whether L and R are in consecutive measures in the music

(Kirlin and Yust 2016, 135-6)

These features are sorted into melodic, harmonic, metrical, and temporal categories as follows

- Melodic: SD-M, SD-LR, Int-LMR, Int-LR, IntI-LR, IntD-LR
- Harmonic: RN-M, RN-LR, HC-M, HC-LR, CT-M, CT-LR
- Metrical: Met-LMR, MetN-LR, MetO-LR
- Temporal: Lev1-LR, Lev2-LR, Lev3-LR

(Kirlin and Yust 2016, 136)

Then these categories are narrowed down and ranked by importance. This yields a hierarchy with harmony at the top, followed by melody, then meter, and finally temporality.

The results showed that harmony is the most important marker for the reductions in terms of harmonic context and identification of non-chord tones which is obvious for an analyst, but it is important to have the computer achieve the same outcome. Melody is the next most important marker, when harmonic context and non-chord tones do not give enough information about scale-degree progression and interval patterns. Following this, meter is applied to anything that is undetermined. Though this procedure seems obvious to the analyst, the hierarchy of steps is the most important part to the computer because it gives the computer a specific order to follow. To reiterate, this has not been fully tested, but it is useful for the understanding the creation of a Machine Learning model.

4.3 Summary

In this chapter, I have shown a few of the recent developments in Machine Learning applied to Music. I have traced the work of Darrel Conklin in particular, since he is a pioneer in the field and continues to contribute to research. As noted above, I have not included a critical analysis section, because the comments I would have made there are more appropriate to the concluding chapter of the thesis, since they address the current state of the field.

Chapter 5- Conclusion

As noted earlier in this thesis, there are different streams in Computer Music Analysis and I have concentrated on Music Information Retrieval (MIR), Optimization, and Machine Learning. These streams often run in parallel because of their different goals. In my concluding chapter I consider some of the most recent developments in Temperley's work, offer methods to bridge the parallels, and present solutions, both general and specific, for the five critical issues mentioned in Chapter 1.

5.1 Further Temperley Research and Probability

Following the *Cognition of Basic Musical Structures* (2001) and *Music and Probability* (2007) Temperley continued his work on borrowing music-like concepts from other disciplines. Two articles, "Information Flow and Repetition in Music" (2014) and "Information Density and Syntactic Repetition" (2015) adapt concepts from other disciplines to further Optimization.

The first article adapts uniform information density as a methodology, which is probability based, borrowed from psycholinguistics and used to further explain parallelism—when parts of a musical work are repeated in an exact or similar fashion and, thus, can be considered as "parallel." Temperley renamed the concept "information flow for repetition in music" and tested it on the Barlow and Morgenstern corpus of musical themes⁸. Temperley found that in parallel sections of a piece the repetition is often more chromatic, but where this is the case the overall piece has a higher probability of smaller diatonic intervals. Thus, the

⁸ A set of 10,000 themes available in print under co-authors Barlow and Morgenstern

juxtaposition of chromatic and diatonic intervals makes the parallelism stand out. Temperley also notes that harmony impacts the repetitions.

The second article looks even closer at parallelism and information flow. It states that “less probable events convey more information” (Temperley and Gildea 2015, 1802). Temperley’s conclusion is consistent with what is referred to in the analysis of prose in the “Inference in an Authorship Problem” (Monstellar and Wallace 1963). This article explains that specific words indicate more than others about an author. I notice that by potentially using Poisson Process and negative binomials—two standard concepts in Probability and Statistics—the specific author of a passage in a multi-author work can be found. This links to Temperley because they follow the acceptance of Bayes’ Rule and are, therefore, part of Bayesian Probability. Temperley’s most recent contributions to the field of Computer Music Analysis is this multi-disciplinary borrowing of research tools. It is the interdisciplinary approach more than any other development that holds greatest potential for the field

5.2 Machine Learning as a means to an end

Machine Learning concentrates on the tool itself. Since this is the most recent development in computer research, and touches on Artificial Analysis, I have left the critique until the conclusion. Because Machine Learning focuses on the tool, it does not have a larger goal other than creating a better tool. This method is best used, in the grand scheme of Computer Music Analysis, as a way to improve and bring other aspects of Computer Music Analysis closer to its goals.

5.3 CompMusic as an example of Intersection

Some methods of using different streams of Computer Music Analysis have been suggested by the authors cited throughout the thesis, but I would like to add my own suggestion: researchers need to coordinate more closely in developing their work. I believe this will further the goals of MIR, Optimization, and Machine Learning. I will focus on “CompMusic,” since it brings together several previously unconnected avenues of research. In this regard, it can serve as an example for the rest of the Computer Music Analysis community to emulate.

CompMusic, also known as *Computational Models for the Discovery of the World’s Music* aims to investigate non-western music. More specifically, “its goal is to advance music analysis and description research through focusing on the music of specific non-Western musical cultures” (CompMusic Project and Workshops 2012, 8). The research project is supported by the European Research Council and the coordinator, Xavier Serra is centred in Spain (CompMusic Website <http://compmusic.upf.edu/>). CompMusic has used multiple streams to finish their database within a few years—2011 to 2017. It seeks “to challenge the current Western centered information paradigms” (*CompMusic*). It concentrates on five traditions of world music: Hindustani, Carnatic, Turkish-makam, Arab-Andalusian, and Beijing Opera (*CompMusic*). Music research has traditionally focused on Western Music, so the researchers for CompMusic had to start from very little. Because of their short time frame, probability, statistical models, and machine learning were used.

Within CompMusic, Machine Learning is used to solve specific problems that hinder the progress of the database, such as in the structure analysis of Beijing Opera (Yang 2016). Initially, resources such as probabilistic and statistical models were used to find novel ways to solve specific problems. For example, with Maghreb, a Moroccan type of music (which is a subset of

Arab-Andalusian music), annotation was difficult, so a tool was created to fix these issues (Sordo et al. 2014). These methods were then adapted to be used in the database.

Since combining different approaches in Computer Music Analysis worked well for CompMusic, I can foresee that the same could work for an MIR project like SIMSSA. To me, it appears that researchers are not sharing their tools and procedures to an optimal degree. This is partially due to a geographic issue, since researchers in MIR, Optimization, and Machine Learning seem to be in different parts of the world. If David Temperley, Darrell Conklin, and members of the SIMSSA project, such as Ichiro Fujinaga and Andrew Hankinson, were to share their tools and approaches more closely, I believe that there could be many new creative problem-solving methods. One example is the previously mentioned solution to the authorship problem (mentioned in Further Temperley research).

5.4 Five general areas for improvement in the field

In writing this thesis, I have observed five general areas where improvement can be made. What is needed is the following: first, an institutional critical analysis of the field; secondly, a closer coordination between Optimization and Machine Learning; thirdly, research into authorship; fourthly, exploration into new areas in Machine Learning; and lastly, closer integration of various MIR resources in developing Optimization and Machine Learning.

1. Critical analysis in Computer Music Analysis as a distinct enterprise has not been performed up to this point except for MIREX, Music Information Retrieval Evaluation eXchange.

MIREX, in brief, is a “framework for the formal evaluation of Music Information Retrieval (MIR) systems and algorithms” (Downie 2008, 247). The goal of MIREX is to investigate the specific tools and algorithms that are the building blocks of larger databases. This method isolates approaches that are nearing the end of their life cycle and compares the performance of systems with similar goals. This provides data about the accuracy and projected utility of algorithms to researchers who want to work within MIR. MIREX, however, only looks at MIR tools and concentrates heavily on methods examining audio data. It does not seem to consider specific issues and how they can be solved using other streams in Computer Music Analysis. Presumably this limitation will be overcome in the future.

2. Closer coordination between Optimization and Machine Learning.

Optimization and Machine Learning have different goals. Optimization, as I have defined it, aims to use computers to mimic a human perception in music to understand the brain. Machine Learning wants to create the specific tool to complete a specific task. However, the end products created by both streams can be used to solve specific problems and tasks in MIR as shown by CompMusic.

3. Research into authorship.

In terms of specific items for research, the areas of authorship and what makes a piece a composer’s own work, has room for growth. This is important for proper identification of a work’s author when it is unknown. This is a common problem with ancient music. Fresh

research could involve the methods put forth by Monstellar and Wallace in 1963 with recent Cathé research on harmonic vectors and their uniqueness to the composer (Cathé 2010a), and Temperley's research on information flow (Temperley 2014), Bayesian Probability (Temperley 2007), and Syntactic Repetition (Temperley and Gildea 2015).

4. New areas in Machine Learning.

Machine Learning has concentrated on music generation and, by using probabilistic and statistical analysis, the music generation can improve by keeping high probability events. Machine Learning can also branch out into more analytical pursuits by mean of analytical algorithms used in Optimization to 'teach' a computer to do analysis. This could improve the current analysis available in Optimization and help to further mimic human perception in the machine.

5. Closer integration of various MIR resources in developing Optimization and Machine Learning.

I have offered specific examples of Optimization and Machine Learning aiding in the creation of an MIR database. However, the opposite development could occur, where MIR databases could be used to develop new research tools. In particular, Humdrum, an analytical MIR tool, has a reserve of files that can be used for both Machine Learning and Optimization. Similarly, various corpora of music assembled in MIR databases could be used as test sets for the same purpose.

5.5 Persaud's Five Critical Issues with Solutions

This thesis has begun the task of a critical analysis by showing different tools in Computer Music Analysis as a whole. The tools selected are of different ages, sizes, and have different researchers associated with them, but all aim to use the computer as their means to an end. I shall conclude the thesis by returning to a set of five particularly acute problems in the field, which I mentioned in my introduction

1. Human error: The problem of human error can be resolved by the creation of more accurate algorithms—either by using harmonic vectors or one of the many Temperley models
2. Specifying input: Improvement in specifying input are imperative to the growth of the field. A researcher reading articles or using pre-existing model needs to know what input should be used. This can be fixed by specifying the input in greater detail in articles and by creating genre-specific standards.
3. Consistent evaluation principles: It is necessary to extend principles used for MIREX to other branches of Computer Music Analysis. Overall, more critical work needs to be done in Computer Music Analysis. Having principles or guidelines will assist in this venture.
4. The interdisciplinary problem of a Lingua franca: To solve this problem Computer Music Analysis should create universal or at least common standards and modes of discourse for describing computer research in music. There are standards for MIR in terms of research tools, algorithms, and systems but those researchers not working in the area are not aware of them. And because many of the tools and procedures are borrowed from other areas of computer research, they are applied in different ways in specialized music research.

5. “What’s the Point?”-Undefined goals. The broader audience needs to understand why Computer Music Analysis is important. This can be overcome by looking at the broader scope of each branch.

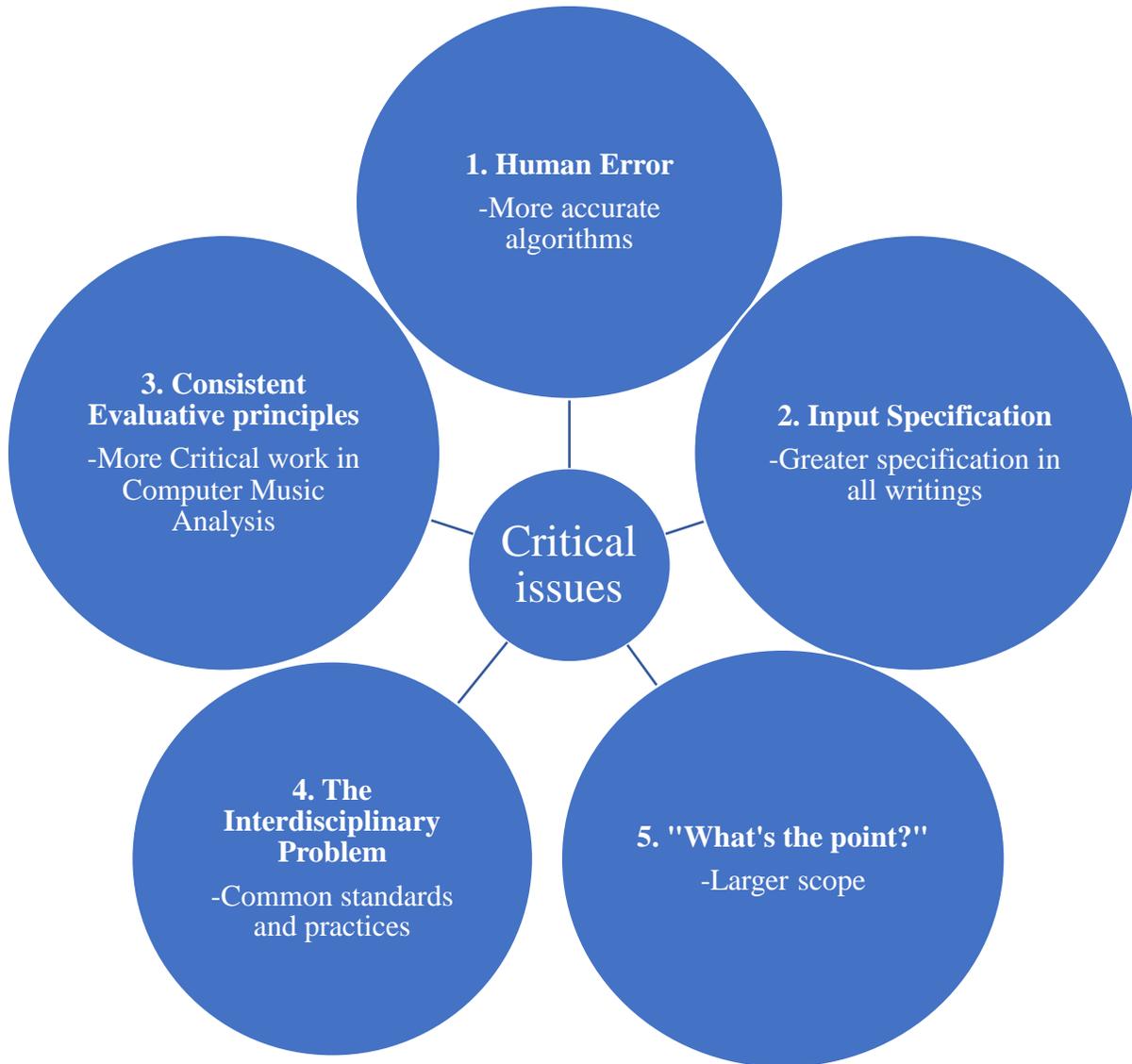


Figure 5 Graphic of 5 critical issues with solutions

In the end, there are multiple avenues to take when it comes to solving the Critical Issues in Computer Music Analysis. Here, I have briefly given my own solutions to these issues and other

aspects and direction for further research, but I have not explained the importance of Computer Music Analysis.

Computer Music Analysis is vital to analysis as a whole because it often adds a quantitative aspect and takes advantage of technology. By incorporating probability and statistics and computational algorithms, the output of the analysis can rely on a mathematical explanation for a qualitative phenomenon. Technology is a fast-growing field and its use in music analysis is inevitable. These new software and hardware move from day to day use into research and improve the field. However, like all changes, it has its own limitations and critical issues. Thesis limits and problems is what fascinates me for this thesis. My overall conclusion is that researchers need to take a critical stance on the discipline for it to grow quickly and efficiently and is a necessity to further improve music analysis.

Bibliography

- Alphonse, Bo H. 1980. "Music Analysis by Computer: A Field for Theory Formation," *Computer Music Journal* 4, no. 2: 26-35.
- Antila, Christopher, Julie Cumming et al. 2014. "Electronic Locator of Vertical Interval Successions. *Montreal Digital Humanities Showcase UQAM*. (Available as slides, scripts, and poster via Elvis website)
- Appleton, Jon. 1986. Review of *Composers and the Computer* by Curtis Roads. *Musical Quarterly* 72: 124.
- Birmingham, William, Roger Dannenberg, and Bryan Pardo. 2006. "Query by Humming with the Vocal Search System." *Communications of the ACM* 49, no.8: 49-52.
- Bozurt, Barış and Karaçali Bile. 2015. "A Computational Analysis of Turkish Makam Music Based on a Probabilistic Characterization of Segmental Phrase," *Journal of Mathematics and Music* 9, no. 1: 1-22.
- Burgoyne, John Ashley, Ichiro Fujinaga and J. Stephen Downie. 2016. "Music Information Retrieval." In *A New Companion to Digital Humanities* edited by Susan Schriebman, Ray Siemens and John Unsworth, 213-228. Wiley.
- Cambouropoulos, Emiliós. 2006. "Musical Parallelism and Melodic Segmentation: A Computational Approach," *Music Perception: An Interdisciplinary Journal* 23, no 3: 249-268.
- Cantus Ultimus. *SIMSSA*. <<https://cantus.simssa.ca/>>
- Cathé, Philippe. 2010a. "Harmonic Vectors and Stylistic Analysis: a Computer-aided Analysis of the First Movement of Brahms' String Quartet op 51-1," *Journal of Mathematics and Music* 4, no 2: 107-119.
- . 2010b. "Nouveaux Concepts et Nouveaux Outils pour les Vecteurs Harmoniques" *Musurgia* 17 no 4: 57-79.
- "CompMusic Project and Workshops." 2012. *Computer Music Journal* 36, no. 4: 8.
- CompMusic*. Music Technology Group, n.d. Web. Accessed 04 Mar. 2017.
<<http://compmusic.upf.edu/>>
- Computer Music Journal*. MIT Press Journals.
<www.mitpressjournals.org/cmj >
- Conklin, Darrell. 2006. "Melodic Analysis with Segment Classes." *Mach Learn* no 65: 349-360.
- . 2008. "Discovery of Distinctive Patterns in Music." *International Workshop on Machine Learning and Music*.

- , 2010. "Distinctive Patterns in the First Movement of Brahms' String Quartet in C Minor," *Journal of Mathematics and Music* 4, no. 2: 85-92.
- , 2016. "Chord Sequence Generation with Semiotic Pattern," *Journal of Mathematics and Music* 10, no 2: 92-106.
- Conklin, Darrell and Ian H. Witten. 1995. "Multiple Viewpoint Systems for Music Prediction," *Journal of New Music Research* 24, no 1: 51-73.
- Cuthbert, Michael Scott. "Music21: A Toolkit for Computer-Aided Musicology." *Music21: A Toolkit for Computer-Aided Musicology*. N.p., n.d. Web. 07 Mar. 2017.
- Dannenber, Roger B. 2007. "A Comparative Evaluation of Search Techniques for Query-by-Humming Using the Musart Testbed." *Journal of the American Society for Information Science and Technology* 58, no 5: 687-701.
- De Haas, W. Bas et al. 2013. "Automatic Functional Harmonic Analysis," *Computer Music Journal* 37, no 4: 37-53.
- Desain, Peter and Henkjan Honing. 1992. "Time Functions Function Best as Functions of Multiple Times." *Computer Music Journal* 16, no 2: 17-34.
- , 1999. "Computational Models of Beat Induction: The Rule Based Approach" *Journal of New Music Research* 28, no 1: 29-42.
- Donnelly, Patrick J. and John W. Sheppard. "Classification of Musical Timbre Using Bayesian Networks," *Computer Music Journal* 37, no. 4: 70-86.
- Downie, J. Stephen. 2003. "Music Information Retrieval." *Annual Review of Information Science and Technology* 37: 295-340.
- Downie, J. Stephen 2003. "The music information retrieval evaluation exchange (2005-2007): A window into music information retrieval research," *Acoustical Science & Technology* 29, no. 4: 247-255.
- El-Shimy, Dalia and Jeremy R. Cooperstock. 2016. "User-Driven Techniques for the Design and Evaluation of New Musical Interfaces," *Computer Music Journal* 40, no 2: 35-46.
- ELVIS Project: Music Research with Computers. < <https://elvisproject.ca/>>
- Fujinaga, Ichiro and Susan Forscher Weiss. 2004. "Music" In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens and John Unsworth. Oxford: Blackwell.
- Giraldo, Sergio and Rafael Ramírez. 2016. "A Machine Learning Approach to Ornamentation Modelling and Synthesis in Jazz Guitar," *Journal of Mathematics and Music* 10, no 2: 107-126.
- Giraud, Mathieu et al. 2015. "Computational Fugue Analysis," *Computer Music Journal* 39, no 2 : 77-96.

- Gulati, Sankalp et al. "Time- delayed melody surfaces for raga recognition." *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR'16)*, New York City (USA).
- Hankinson, Andrew, Evan Magoni, and Ichiro Fujinaga. "Decentralized Music Document Image Searching with Optical Music Recognition and the International Image Operability Framework." In *Proceedings of the Digital Library Federation Forum*. Vancouver, BC, 2015
- Hardestry, Jay. 2016. "A Self-Similar Map of Rhythmic Components," *Journal of Mathematics and Music* 10, no. 1: 36-58.
- Helsen, Kate et al. 2014 "Optical Music Recognition and Manuscript Chant sources," *Early Music* no. 42: 555–58.
- Huron, David. 1988. "Error Categories, Detection and Reduction in a Musical Database," *Computers and the Humanities* no. 22: 253-264.
- , 2001. "Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles," *The Journal of the Acoustical Society of America* 19, no 1: 1-64.
- The Humdrum Toolkit: Software for Music Research. 2001.
<http://www.musiccog.ohio-state.edu/Humdrum/FAQ.html>
- Iñesta, José M., Darrell Conklin, and Rafael Ramírez. 2016. "Machine Learning and Music Generation," *Journal of Mathematics and Music* 10, no. 2 :87-91.
- Kacprzyk, Janusz and W. Ras Zbigniew. 2010. *Advances in Music Information Retrieval*. Berlin: Springer International Publishing
- Karasmanoğlu, M. Kemal. 2012 "A Turkish Makam music Symbolic Database for Music Information Retrieval: SymbTr," *Proceedings of ISMIR*.
- Keller, Robert et al. 2013. "Automating the Explanation of Jazz Chord Progressions Using Idiomatic Analysis." *Computer Music Journal* 37, no.4: 54-69.
- Kirlin, Phillip B, and Jason Yust. 2016. "Analysis of Analysis: Using Machine learning to Evaluate the Importance of Music Parameters for Schenkerian Analysis," *Journal of Mathematics and Music* 10, no. 2: 127-148.
- Larson, Steve. 2004. "Musical Forces and Melodic Expectation: Comparing Computer Models and Experimental Results," *Music Perception: An Interdisciplinary Journal* 21, no. 4: 457-498.
- Louridas, Panos and Christof Ebert. 2016. "Machine Learning," *IEEE Software*: 110-115.
- Manning, Peter et al. 2001. "Computers and Music." *Grove Music Online*. Oxford Music Online. Oxford University Press, accessed April 11,2017
 <<http://www.oxfordmusiconline.com.proxy.bib.uottawa.ca/subscriber/article/grove/music/40583>>

- Meeus, Nicolas. 2003. "Vecteurs harmoniques" *Musurgia* 10, no 3: 7-34.
- Meredith, David. 2016. *Computational Music Analysis*. Springer Cham, Heidelberg, New York, Dordrecht, and London: Springer International Publishing Switzerland.
- Monsteller, F and David L Wallace. 1963 "Inference in an Authorship Problem," *Journal of the American Statistical Association* 58, no. 302: 275-309.
- Music21: A Toolkit for Computer-Aided Musicology. < <http://web.mit.edu/music21/>>
- Orio Nicola. 2008 "Music Indexing and Retrieval for Multimedia Digital Libraries." In Agosti M. (eds) *Information Access through Search Engines and Digital Libraries. The Information Retrieval Series*, vol 22. Springer, Berlin, Heidelberg
- Pardo, Bryan. 2008. "Music Information Retrieval," *Communications of the ACM* 49, no. 8 : 29.
- Pardo, Bryan et al. 2008. "The VocalSearch Music Search Engine," *JCDL*.
- Patrick, Howard P. 1974. "A Computer Study of a Suspension-Formation in the Masses of Josquin Desprez," *Computers and the Humanities* 8: 321-331.
- Piantadosi, Steven T. et al. 2011. "Word Lengths Are Optimized for Efficient Communication," *Proceedings of the National Academy of Science of the United States of America* 108, no. 9: 3526-3529.
- Ponce de León, Pedro J. et al. 2016. "Data-Based Melody Generation through Multi-Objective Evolutionary Computation," *Journal of Mathematics and Music* 10, no. 2: 173-192
- Roads, Curtis et al. 1986. "Symposium on Computer Music Composition," *Computer Music Journal* 10, no 1: 40-63.
- Search the Liber Usualis. *SIMSSA*. <liber.simssa.ca>
- SIMSSA-Single Interface for Music Score Searching and Analysis. <simssa.ca>
- Smith, Leigh M. and Henkjan Honing. 2008. "Time- Frequency Representation of Musical Rhythm by Continuous Wavelets," *Journal of Mathematics and Music* 1, no. 2: 81-97
- Sordo, Mohamed et al. 2014. "Creating Corpora for Computational Research in Arab-Andalusian Music" *Proceeding of the 1st International Digital Libraries for Musicology workshop* London (UK). < <http://mtg.upf.edu/node/3028>>
- Temperley, David. 2001. *The Cognition of Basic Musical Structures*. Cambridge, Massachusetts and London, England: MIT Press.
- , 2007. *Music and Probability*. Cambridge: MIT Press.
- , 2010. "Modelling Common Practice Rhythm," *Music Perception: An Interdisciplinary Journal* 27, no. 5: 355-376.

- , 2014. "Information Flow and Repetition in Music," *Journal of Music Theory* 58, no. 2: 155-178.
- Temperley, David and Christopher Bartlette. 2002. "Parallelism as a Factor in Metrical Analysis," *Music Perception: An Interdisciplinary Journal* 20, no. 2: 117-149.
- Temperley, David and Danial Gildea. 2015. "Information density and Syntactic Repetition," *Cognitive Science*, no. 139: 1802-1823.
- Tenkanen, Atte. 2010. "Tonal Trends and α -Motif in the First Movement of Brahms' String Quartet op. 50 mvt. 1," *Journal of Mathematics and Music* 4, no. 2: 93-106.
- Viglilensoni, Gabriel et al. 2011. "Automatic Pitch Recognition in Printed Square-Note Notation." *Proceedings of 12th International Society for Music Information Retrieval Conference* Miami, Florida: 423-428.
- Wang, Ge, Perry R. Cook, and Spencer Salazar. 2015. "ChucK: a Strongly Timed Computer Music Language," *Computer Music Journal* 29, no. 4: 10-29.
- Yang, Yile. 2016. "Structure Analysis of Beijing Opera Arias" Master Thesis, Universitat Pompeu Fabra, Barcelona (Spain).