HOKKAIDO UNIVERSITY

Title	和声学規則と複数アルゴリズムを応用した自動伴奏生成システムの提案と実装
Author(s)	エバンズ, ベンジャミン ルカ
Degree Grantor	北海道大学
Degree Name	修士(情報科学)
Issue Date	2015-03-25
Doc URL	https://hdl.handle.net/2115/58266
Type	master thesis
File Information	Benjamin-MasterThesis.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

修士論文

和声学規則と複数アルゴリズムを応用した自動伴奏生成システムの提案と実装

北海道大学 大学院情報科学研究科
複合情報学専攻 複雑系工学講座 ヒューマンコンピュータインタラクション研究室

エバンズ　ベンジャミン　ルカ

2015 年 2 月 10 日

Master's Thesis

Implementation of an Automatic Accompaniment Generator using the Theory of Harmony and an Application of Multiple Algorithms

by

Benjamin Luke Evans

Laboratory of Human Computer Interaction,
Division of Synergetic Information Engineering,
Graduate School of Information Science and Technology
Hokkaido University
10th February, 2015

Abstract– In this paper, we describe *"Auto Chorus Creator (ACC)"*, an automatic four-part accompaniment generator with three novel characteristics.

Firstly, we have incorporated the method of Dynamic Programming in the search process of *ACC*, enabling the real-time generation of even long pieces of music. Secondly, we have applied the rules of harmony as continuous functions, as opposed to the binary function approaches many related works have used, allowing the system to compare multiple songs in further detail than previous works and make a more objective distinction between them. Finally, we have also implemented a rule base of heuristic rules not directly related to rules of musicology to use in generating the initial point from which *ACC* starts its search for the optimal output. We have seen this significantly effects the evaluation values of the final output as to when the search is started from a randomly generated point in the search field.

Our research contributes not only to the field of automatic accompaniment generation, but also to the field of automatic composition in general, offering a novel approach of combining multiple algorithms to produce numerous equally-optimal musical solutions in real-time.

Keywords: Automatic Music Generation, Automatic Choral Harmonisation, Theory of Harmony, Dynamic Programming, Heuristic Rule Application,

概 要: 本研究では，さまざまな予備実験を踏まえ，実用的な自動音楽生成アプリケーションを実装するにあたり，システムに求められる条件を 3 つ述べている．また，それらを同時に実現するシステム機構を提案し，自動伴奏生成システム *"Auto Chorus Creator (ACC)"* として実装し，その有用性を検証している．

実用的な自動音楽生成システムに求められる特徴として，本研究では特に以下の 3 つを挙げている．1 つ目に，システムは長い楽曲も実時間内に生成可能でなければならない．本稿ではダイナミック・プログラミングの手法を利用することにより，システムの実時間実行を実現している．2 つ目の特徴として，自動音楽生成システムには，何かしらの形で音楽学の知識が実装されている必要がある．本研究では先行研究が扱ってきた和声学の知識をシステムの評価関数に実装している．その際，楽曲により明確で客観的な優劣差をつけるために，和声学の諸規則を連続関数として再定義し，実装している．自動音楽生成システムに求められる 3 つ目の特徴として，本研究では同じ入力から同等に優れている複数の出力を生成可能であることを挙げている．この実現のために，本研究では出力の探索を行う初期解の生成にヒューリスティック・ルールベースの確率的適用を用い，複数アルゴリズムを組み合わせたハイブリッド形式の作曲機構を提案している．本研究では，上のような実装により，単一アルゴリズムを用いた場合と比べて，統計的に優れた結果が得られていることを紹介している．

本研究は，多数の予備実験，提案アルゴリズム，及び実装システムの紹介と解析により，自動伴奏生成研究の他，自動作曲研究一般にも貢献している．特に，音楽学的に同等に優れている複数解の実時間生成を可能とする本研究の提案は，既存システムの性能向上にも役立つものであると考えている．

キーワード: 自動音楽生成, 四声体和声, 和声学, ダイナミック・プログラミング, ヒューリスティックルールベースの確率的適用

目 次

第 1 章　はじめに **1**
1.1　本研究の概要 1
1.2　本研究の位置づけ 2
1.3　本稿の構成 2

第 2 章　研究背景と予備実験 **4**
2.1　計算機による自動作曲 - 音楽情報処理の研究 4
2.1.1　先行研究 4
2.1.2　作曲における探索領域 6
2.2　和声操作としての作曲 - 音楽学の研究 8
2.2.1　和声の歴史 8
2.2.2　先行研究 9
2.2.3　自動和声付与システムの構築 10
2.3　印象誘導としての作曲 - 生体心理学の研究 14
2.3.1　先行研究 14
2.3.2　作曲による印象誘導実験 15
2.4　自動作曲研究の課題の総括 21

第 3 章　実装システム **23**
3.1　概要 23
3.2　設計詳細 27
3.2.1　ヒューリスティック・ルールベースによる初期解の生成 27
3.2.2　解空間における探索と評価 29
3.2.3　ユーザ視点でのシステム概要 35

第 4 章　システム評価と結果 **37**
4.1　モジュールごとの評価と考察 37
4.2　システムに対する総合的な考察 44

第 5 章　おわりに **46**
5.1　まとめ 46
5.2　今後の展望 46
5.3　謝辞 47

第1章　はじめに

本章では，まず本研究の概要を述べる．その後，本研究の特徴を簡単に説明し，それが自動音楽生成の研究分野においてどのように位置づけられるかを説明する．最後に，本稿の構成を述べる．

1.1　本研究の概要

西洋音楽は，グレゴリオ聖歌等の教会音楽から生まれてきたと言われている．当初は聖職者しか口にすることができなかった「音楽」は，社会の上流階級から徐々に一般の人々にも広がり，楽器や楽譜の発展に支えられ，芸術の一種として次第に西洋社会に広がっていった．いまや「音楽」は，幼児さえ口ずさむほど，私たちの身の回りに完全に浸透し，生活に欠かせない一部となっている．音楽を聴くのみだった我々は，安価な楽器の開発をきっかけに，自身で演奏することに挑戦した．演奏に慣れた我々は，インターネット技術の発展等に起因して情報発信の機会が増えたことをきっかけに，自身の演奏を他者にも聴いてもらうようになった．そして「凡人」の中にも，大衆に受け入れられるような優れた演奏家が生まれると，我々は自身で作曲をするようにまでなった．今では，完成した音源を嗜んだり，与えられた楽譜を自身で演奏・再現するだけでなく，音の並びの元となる何かしらの意図の創造から，それを音として時系列上に複数並べて組み合わせる作曲の過程までを，我々ユーザが自ら取り組むようになったのだ．

ユーザは自身で作曲をするようになったわけだが，それは決して簡単な作業ではない．ランダムに並べた音を音楽と呼ぶこともあるが，それは一般聴衆が無意識のうちに培ってきた，それまでの個人的な音楽聴取に基づく音楽解釈の認知活動では理解し難いものである場合が多い．そのため，そのような乱数的音楽には，それ自身の新たなジャンルが作られ，新たな聴取スタイルが作られてきた．広く一般的な理解を得，更に聴衆の興味を惹きつけるためには，一般聴衆に存在するその無意識的な音楽解釈機構に則った音楽理解や作曲技術が必要となる．つまり，従来の音楽様式の枠組みの中で作曲を行う際には，多面的な前提知識と洗練された作曲技術とを併せ持った上で，それらを複合的に利用することが必要となるのである．気楽に作曲を始めたユーザも，優れた楽曲の創作を行うにあたり，求められるこれらの技術の複雑さを知った時に，自身の作曲がいずれ陥る限界を見抜き，諦めることさえある．

そこで本研究では，これら専門的な技術を持たないユーザの作曲活動をサポートし，いずれ自分で作曲ができるようになるまでの過程をアシストするシステムを構築することをモチベーションとした．一般に自動作曲の研究は複雑で，それを扱った先行研究は多くの分野に渡って複数存在する．しかし，それぞれの分野で得られた知見を包括的に実装しようとした試みは少ない．そこで本研究では，複雑である自動作曲という問題を多面的に捉

えるために，自動作曲研究を扱ったこれら複数の研究分野の先行研究を調べ，予備実験を多数重ねたうえで，最終的なシステムの提案と構築に至った．本研究で作成した自動伴奏生成システムは，従来の研究では利用されていない，複数のアルゴリズムを融合したハイブリッドな作曲機構を特徴とし，ユーザからの同じ入力に対しても，同等に優れた複数の解を実時間内に生成することに成功している．著者は自動作曲に見受けられる一般的な課題をまとめ，本研究で構築したシステムがその一部を解決できることを示した．また，まだ解決できない課題については，今後どのような仕組みがシステムに必要になるのかを考えた．

なお，一般に「音楽」というと，その指し示す対象が各個人によって大きく異なるという実態が見受けられる．また，ある定義の枠組みの中で音楽の様々なジャンルを比べた際にも，その特徴が実に多様で，ある一つの研究でそれらをすべて網羅する，一意的な議論やシステム構築が不可能であると考える．そこで本研究で扱う「音楽」は，古典的なクラシック音楽をはじめ，そこから派生した，一般的な和声理論に則って分析・解釈できる音楽に留めることとする．これには，古典的なジャズや一部のポピュラー音楽等も含まれるが，これらの名称に分類される全ての音楽を扱っているわけではない．本稿内で音楽に言及する部分には，楽譜での表記等を提示しながら具体的な例をなるべく多く挙げているが，紙面の関係上，全ての例を網羅することはできなかった．本稿内で提示している楽譜は鍵盤を叩く等して，実際に聴きながら読み進めてもらえると良いだろう．

1.2　本研究の位置づけ

一般に自動作曲の研究は，1959 年に「イリアック組曲」[1] が計算機による最初の自動作曲組曲として発表されてから，数十年の歴史を経た今も，なお盛んに続けられている．本稿では，多方面に渡る自動作曲の先行研究をまとめ，それぞれの分野において著者が行った予備実験を紹介している．その中で，既存の自動作曲システムではまだ解決できていないいくつかの課題を取り上げ，それらをまとめた．更に，それら課題を具体的に扱うためのシステム・アプローチを提案し，自動伴奏生成システムとしてそれらの機構を実際に実装し，検証した．

本研究は，一般的な自動作曲研究における総合的な課題を複数指摘し，それらを同時に取り扱った研究成果を上げている．従来の自動作曲研究は，広い研究分野の中でも限られた課題を扱い，限定された研究環境での成果は上げながらも，実際のアプリケーションに向けた実践的な実装を行うことができていないものが多い．本研究は自動作曲の複数の側面を総括的に捉えることを試み，特に自動伴奏生成という作曲分野における新たなシステム設計を提案し，続く研究のシステム構築にも役立つ指針と実装設計のプラットフォームを提供している．

1.3　本稿の構成

第 2 章ではまず，様々な研究分野における自動作曲の先行研究を，その分野ごとにまとめて紹介する．また，本研究で行った各分野における様々な予備実験を紹介し，そこで得られた知見をまとめる．その後，従来研究で指摘され，未だその解決に至っていない自動

作曲研究の課題をまとめ，今後の研究で着目する必要のあるシステム設計のポイントを体系的に提案する．

続いて第 3 章では，第 2 章で示した自動作曲研究の課題を踏まえ，本研究で構築した自動伴奏生成システムを紹介する．システムは大きく 2 つの部分から成り立つため，それらを個別に説明した後に，総合的に望まれるシステムの挙動やユーザ視点でのシステム概要等をまとめる．

第 4 章では，実装したシステムについて行った評価実験を説明する．自動作曲研究では，その扱う対象（楽曲）一つ一つが芸術作品として存在し，客観的な評価指標を割り当てることが非常に困難である．そのため，先行研究における楽曲評価の多くは主観的な評価プロセスに頼る傾向がある．本研究では，新しく定義した評価関数の特性により，システムが生成する楽曲をより客観的に評価することに成功している．本章ではシステム内の各部分単体での評価，及びシステムの総合的な評価の結果を紹介し，それらについての考察を行う．

最後に第 5 章では，本研究の概要を再度述べ，成果をまとめ，今後の展望を記載する．また，本稿末尾には参考文献一覧の他に，著者の発表論文一覧も掲載しているので，必要に応じて参照していただきたい．

第2章　研究背景と予備実験

自動作曲に関する研究は，様々な視点から行われてきた．本章では，それらを音楽情報処理分野の研究，音楽学分野の研究，生体心理学分野の研究に大別し，それぞれの観点ごとに先行研究を紹介する．また本研究では，自動伴奏生成システムを構築する際，各分野における予備実験も複数行っているため，それらの詳細と，そこで得られた成果とを紹介する．最後に，それら予備実験から得られた知見をもとに，自動作曲研究における様々な課題をまとめ，それらに対処した研究の指針を提案する．

2.1　計算機による自動作曲 - 音楽情報処理の研究

情報処理研究界には，音楽と情報を専門的に扱う研究分野「音楽情報処理」がある．その研究範囲は自動作編曲，自動演奏，歌声合成，楽曲検索等，多岐に渡る．本節では，特に自動作編曲分野の先行研究を紹介し，様々なシステムに実装されてきた多くのアルゴリズムと，それら先行研究の特色をまとめる．またその後，作曲における解の個数の理論値を計算し，作曲における解の探索領域が指数的に増大することを示す．この解空間の特性より，一般的に自動作曲システムは効率的な解探索アルゴリズムを必要とし，全探索では問題の解決が難しいことを説明する．

2.1.1　先行研究

膨大な解空間から，出力する楽曲を効率よく選ぶために，自動作曲に関する先行研究では様々なアルゴリズムが利用されてきた．そのシステムの形式には，人間が作曲する際の認知活動を模倣したようなものもあれば，マシンパワーを生かした計算機特有のものもある．以下に，それらを簡単に紹介する．

自動作曲研究の中で最も代表的な手法として，作曲プロセスを 制約充足問題 として定式化したものが挙げられる．古典的なクラシック様式の楽曲を創作する際，制約は単旋律の入力楽譜やコード進行列等で与えられることがある [2], [3]．システムはその後，和音や拍子の決定に音楽学の規則体系を当てはめ，最適解を求める．他にも，入力に歌詞を使ったシステムもある [4]．深山らは和語に存在する言葉の韻律を自動抽出し，それを旋律の動きにマッピングしたシステム *Orpheus* を構築し，オンライン上で公開している [1]．既知の音楽学知識（和声学等）は規則群の形で書かれることが多く，計算機に実装しやすいという特徴がある．そのため，制約充足問題への定式化は比較的単純なプログラム実装で済む場合が多い．

[1] 自動作曲システム *Orpheus* `http://www.orpheus-music.org/v3/`

和声などのように，体系化が比較的完結している楽曲要素もあれば，リズムの制約などのように，今現在も規則の具体的な形式化に乏しい楽曲要素もある．楽曲に内在する，体系的にまとめられていないそのような知識をうまく抽出し，モデルとして構築する方法として**機械学習**の手法が利用されている．また，その楽曲要素の情報をわざわざ抽出することなく、既存楽曲から部分的に切り出した音列などをうまくつなぎ合わせて新しい楽曲を作り出す**パターンマッチング**などの手法もある．先に言及した *Orpheus* は，旋律の決定は制約充足的に行うものの，リズムについてはシステム側で用意したリズムパターンから選択して利用するようになっている [4]．他にも，Cope のシステムは機械学習によるパターン認識を利用し，完成度の高い楽曲を生成したことで有名である [5]．Cope は特定の作曲家（モーツァルト等）の既存の楽曲を多数入力し，旋律の進行パターン等を計算機に学習させることによって，その作曲者の作曲スタイルを模倣した楽曲の生成を試みた．

他の自動作曲手法として，**確率モデル**によるアプローチがある．Hild らは，入力の楽譜に対する和声付けをニューラルネットワークを用いて行った [1]．古典派クラシック曲の作曲者バッハの様式に則った和声付けについては良い成果が出ているが，Hild らもこの手法の一般的な作曲への拡張は難しいと述べている．他の確率的手法として，マルコフモデルを用いたシステムもある．マルコフ過程は，限定された記憶に基づき，意味のあるシーケンスを生成することができると仮定されているモデルである．しかし，時間軸上に大域的な関係（フレーズや旋律の係り結びなど）を内在すると知られている音楽に適応する際には，それが確率モデルにうまく反映できない難点がある．Pachet らは制約を設けたマルコフモデルで，その大域的な関係を模倣できる自動作曲機構を提案している [6]．

Phon-Amnuaisuk らは，**遺伝的アルゴリズム**を用いた作曲システムと，同じ音楽学知識を実装したルールベースシステムとの性能を比べる実験を行った [7]．彼らは，音楽学知識がシステム自身に内蔵されるルールベースシステム（たとえば Ebcioğlu の *CHORALE*[8] 等）は，特定のジャンルや音楽様式に対応した自動作曲には優れているものの，システムの拡張が難しく，汎用性がない等という指摘をしている．自動作曲システムが出力する楽曲の特性は，システムに包括される音楽学の知識に直接影響されるため，柔軟な作曲を行うためには，それら音楽学知識もシステムのアルゴリズムから切り離され，柔軟に拡張・変更可能であることが必要だと述べられている．遺伝的アルゴリズムは，他にもジャズの自動ソロ創作等にも使われている．Biles らはシステム *GenJam* において，システムの演奏に対するリアルタイムの評価をフィードバックとして与え，システムが演奏中に自身の演奏を改善していくようなモデルを考えた [9]．*GenJam* にはコードマッピングや調決定に関するいくつかのルールベースが実装され，容易に拡張できる仕組みが整っている．

更に，MacCallum らは一般聴衆による評価を取り入れた実験システム *DarwinTunes* を開発し，研究している [10]．彼らは 7 千人近い被験者による，ネットで公開される 8 秒の短い楽曲に対する評価を利用し，その評価をもとに選択・交叉・突然変異を実施していく 2,500 世代以上に渡る，大規模な遺伝的アルゴリズムの実験を実施した．その中で，人間が好むと考えられる音響的特徴を抽出したり，好印象を誘発する和声やリズム等の楽曲要素を検討している．

近年は，**音楽情報検索**（Music Information Retrieval : MIR）分野の知見を用いた自動作曲システムも開発されている．Devies らの *AutoMashUpper* は，楽曲検索で利用されていた楽曲構造分析技術を用いて，区切られた各区間における楽曲の和声構造が似ている楽

曲同士を重ね合わせたりし，編曲や演奏を行う [11]．楽曲検索分野から派生した自動作曲システムについては，Humphrey らがまとめている [12]．

以上のそれぞれのアルゴリズムは，四声体和声の自動作曲においても用いられている．つまり，広く一般的な自動作曲システムの中でも，特に，入力される単旋律の楽譜を四声部の合唱譜に書き換える（伴奏を追記する）伴奏生成システムにおいて，この様々な探索手法と和声理論を組み合わせた実装が行われてきているのである．Allan らは，隠れマルコフモデルを用いた機械学習型和声付加システムを実装している [13]．鈴木らも，機械学習により確率モデルを構築するシステムを実装しているが，システム内でコードを表現するノードを導入した場合としなかった場合との実行結果の違いを調べている [14]．Ebcioğlu は，この四声体和声の作曲課題用に新たなプログラミング言語を開発し，問題に特化した解探索手法を提案している [8]．

2.1.2 作曲における探索領域

一般に「自動作曲システム」と呼ばれるものは，大きく二種類に分類される [15]．一つは，与えられた旋律に伴奏を付加するシステムであり，もう一つは，旋律そのものの創作も行うものである．前者は，旋律という入力から和声や拍子における制約を抽出し，その制約の中で伴奏を生成する．後者は与えられる音楽学的制約が極端に少なくなるため，解の探索領域が膨大に増えてしまう．そのため，音楽学には直接関係のない他の制約も導入することにより，効率的な作曲を試みる場合が多い．

一般論として，楽曲の可能性は無数に存在し得る．四声体和声に限定して論じた場合も，理論的に存在し得る楽曲の解の数は膨大だ．探索時間やマシンパワーに制約がないのであれば，全探索を行ってすべての解の評価値を比較することが最も簡単な解決定の方法である．しかし一般に自動作曲アプリケーションは実行時間に大きな制約があり，ほぼリアルタイムでの作曲が求められることが多い．このような実際的な作曲環境を前提にシステムを構築するのであれば，その解空間を効率良く探索する手法が不可欠となる．

以下では，一般的な四声部の合唱譜に様々な制約を設けた場合に，システムが生成可能な楽曲の論理的な数を導出する．これは自動作曲の探索領域の広大さを示し，実時間実行の難しさを一つ確認するために行う．著者が調べた限り，このようにシステムの実行時間を議論した先行研究はない．

なお，計算の大前提として，各パートの音域は『和声–理論と実習–』[16]（通称『芸大和声』．第 2.2.2 節参照）に記載されている図 2.1 のものを利用する事とする．声域は多くの楽器の音域よりも狭いため，ここで既に一般的な自動作曲よりも解空間が大幅に縮小されることになる．

一般に，全てのパートが同時に発音し，各パートの音数 n が等しいとき，作曲可能な楽曲の数は

$$(\text{ソプラノの音域} \times \text{アルトの音域} \times \text{テナーの音域} \times \text{バスの音域})^n \tag{2.1}$$

だけ存在する．この式に図 2.1 の音域を適用すると，

$$(22 \times 20 \times 22 \times 22)^n = 212,960^n \tag{2.2}$$

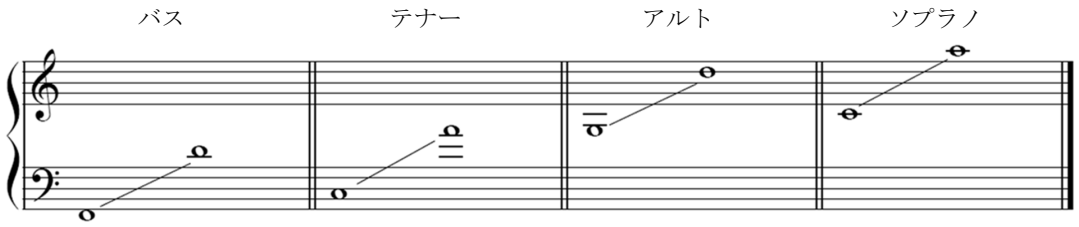

図 2.1: 四声体和声における各声部の基本的な音域

の可能性が存在することになる．これをすべて探索して出力解を決定する（作曲する）ことは現実的ではないので，自動作曲システムはいくつかの制約を設け，その探索を枝切りしなければいけない．もしくは，探索アルゴリズム自体に何かしらの評価関数を導入し，スマートな探索を行わなければいけない．第 2.2.3 節で紹介する，著者が構築した自動和声付与システム *“CMY”* においては，4 パートある内の 1 パートの楽譜を入力することになっているので，その分だけ探索領域が減る．たとえばバス譜を入力した場合，システムが出力可能な楽曲数は

$$(22 \times 20 \times 22 \times 1)^n = 9,680^n \tag{2.3}$$

に抑えられる．しかし，これもまだ実時間実行には探索領域が広大過ぎるため，他の制約も考えなければいけない．

例えば，合唱という音楽形態において，パート同士が交叉[2]することは，実際の作曲ではあまりない．少なくとも,『芸大和声』ではそれを避ける場合が大半である．交叉を許さない場合のバス課題[3]の解の数は，平均すると $3,637.14^n$ に減る．更に,『芸大和声』の第 I 巻にあるように，隣接するパートの音間の距離を制限する[4]こともできる．そうした場合，出力可能な楽曲の平均数は $2,044.14^n$ に減る．

CMY では，探索の領域を更に制限するために，各パートの音は，その時のバス音が指定する三和音（及び一部の四和音）の構成音であることとした．すなわち，バス音が C の音であり，その時の和音が C の和音である場合，残り 3 声部の音も C 和音の構成音 C・E・G のいずれかをとる必要がある．そうした場合，システムが生成し得る楽曲の平均数は 49.45^n 通りに限定される．なお，他の条件下で出力可能な楽曲の理論数は，表 2.1 に示した．

このように，多くの制約を設けた場合でも，出力可能な楽曲の数は膨大で，それは音数 n に対して指数のオーダーで増加する．作曲に際して，音楽学的制約を増やすことにより，その解領域を縮小することは可能であるが，それに応じて，出力される楽曲の特徴は限定されていく．これは，出力する楽曲の新規性等に芸術価値があるとした場合，真に特徴の

[2]和音内に音を配置するにあたって，各声部の正規の音高の順序が入れ替わること（[16] 第 I 巻 122 ページより）．つまり，ソプラノ音がアルト音より低くなったり，バス音がテナー音より高くなったりと，ソプラノ～アルト～テナー～バスという音高の順序が入れ替わること．

[3]和声学の授業において，学習した諸規則の定着と応用を図るために利用される練習問題．単旋律（バスやソプラノが多い）の書かれた楽譜を出題され，そこに和声の禁則に反しないように残り 3 パートを追記する課題.『芸大和声』[16] にも多くの問題が記載され，別巻にはその模範解答も全て掲載されている．

[4]『芸大和声』第 I 巻では，ソプラノ音～アルト音間，アルト音～テナー音間の距離（音高の差）をオクターブ（8 度）まで，テナー音～バス音間の距離は 12 度までとしている．([16] 第 I 巻 18 ページ)

表 2.1: バスパート譜を入力として与えた際に，和声付けシステムが各種条件下で生成し得る楽曲の理論数（n は入力のバス音の数を示す）

		全音			三和音		
		交叉有り	交叉無し	隣接距離制限	交叉有り	交叉無し	隣接距離制限
一般四声体		$212,960^n$	$80,017^n$	$44,971^n$	$3,350^n$	$1,650^n$	$1,088^n$
バス課題	最大	$9,680^n$	$5,180^n$	$2,711^n$	180^n	109^n	66^n
	最小	$9,680^n$	952^n	868^n	100^n	36^n	32^n
	平均	$9,680^n$	$3,637.14^n$	$2,044.14^n$	152.27^n	75.00^n	49.45^n

異なる楽曲を作り出すことが困難になることを意味し，様々な特徴を持った楽曲を同じシステムから出力しようとした場合にも都合が悪くなってしまう．そこで，制約決定の他に，解空間における探索アルゴリズムそのものを工夫することも必要となる．そのために，第 2.1.1 節で紹介したように，多くのアルゴリズムが自動作曲研究で試され，利用されてきているのだ．

2.2 和声操作としての作曲 - 音楽学の研究

音楽は，メロディー，ハーモニー，リズム等といった様々な要素から成り立っている．それぞれに対応する和語として，旋律，和声，拍子等が挙げられる．この中でも，本研究では作編曲の全てのプロセスに関与する [17] と言われている「和声」に着目した．つまり，様々な音楽要素を設計する総合的な作曲行為の中から，まずは楽曲のハーモニーを構築する部分を自動化するシステムを作ることとした．

本節では，まず和声の概念がどのように生まれ，どのように体系化されてきたかを簡単にまとめる．その後，作曲における和声の存在に着目した先行研究を紹介し，その成果をまとめる．最後に，第 3 章以降で説明する本稿の主となる自動伴奏生成システムを実装する前に，予備実験的に構築した自動和声付与システムについて説明する．

2.2.1 和声の歴史

和声学とは，和声を体系的にまとめた音楽学の分野であり，現代の日本における芸術大学や教育大学の音楽専攻の学生は皆，それについて学ぶことになっている．ここで一つ，和声学が日本の伝統音楽（雅楽等）における和音構造等をまとめた「和の国」独特の学問ではないことに注意していただきたい.「和声」は英語「harmony」の訳語であり，ジャンルを限定せず，あらゆる音楽の和音構造を論じる際に使われる一般的な単語である．クラシックにも，ジャズにも，雅楽にも，演歌にも，それぞれ固有の和音体系が存在し，その構造もそれぞれ異なる．しかし便宜のため，以下本稿では「和声」と述べるときに，第 1.1 節で説明した音楽ジャンルに関する和声を指し示すこととする．

現代音楽の起源がどこにあるのか，ということに関しては様々な考え方があるが，少なくとも古代文明にその発祥ルーツがあることは確かである．紀元前数千年から人類は既に

複数の楽器を同時に演奏し，それらを和声的に解釈しながら[5]聴いていた記録がある [18], [19], [20]．音楽一般が形式立てて議論され始めたのはその後，紀元前後のギリシア文明の時代であるとされるが，その頃は旋律の概念や，音楽が人間に与える精神的影響力が主に論じ合われ，和声に関する細かな議論はされていないようだ [18]．

紀元後 9 世紀頃までの西洋音楽は，その詳細について未だに多く議論されているが，その和声体系がオリエント，つまり，シリア，ヘブライ，ビザンチン等からもたらされたとする説が強くなっている [21]．10 世紀から 13 世紀にかけて，楽曲を紙面に書き連ねる記譜法がある程度定着し，現代の記譜法と似たものに変化していった中で，音楽の様々な表層構造[6]が形式立てて議論されていった [22]．特に，ポリフォニー[7]において平進行[8]，反進行[9]，移調[10]等が頻繁に用いられるようになるに連れ，和声の体系化が必需的になっていった [21]．

16 世紀から 17 世紀にかけては，ピアノ譜や合唱譜が現代と同じ大譜表[11]で書かれるようになり，現代使われる和音（3 音以上からなる和音とその転回系）の概念がいよいよ確立された．18 世紀にもなると，いよいよ平均律が音楽界に受け入れられ，広く利用されるようになり，今日の作曲活動の基礎となっている和声学が大成された．この時代の有名な作曲家たち（バッハ，モーツァルト，ベートーベン等）は，今でも多くの人に知られている．本研究に用いた和声学も，この時代のものに基づいてまとめられたものである．

2.2.2 先行研究

江村ら [17] は，音大の作曲家の学生を例に，学術的な作編曲の流れを以下のように述べている．

1. ケーデンス[12]を持つ簡単な和音列をイメージする．
2. 各和音の和声音を考慮しながら適当な律動を持った旋律音を選択していく．
3. 各和音の和声機能を考慮し，代替可能な他の和音に置換する（リハーモナイゼーション）．
4. 各和音にテンション（非和声音）を付加し，和声的意味づけを行う（テンション・ヴォイシング）．

---

[5]つまり，複数の楽器のその様々な響きの組み合せには，それぞれ固有の意味が表現されていると認識していた

[6]表層構造：楽曲を構成する様々な要素の内，音高，音程，拍子等，楽譜に記載可能で，リスナーの耳に直接届くような情報，及び，楽曲名，出版年，アーティスト名等，楽曲に関するメタな情報のこと．深層構造：作曲者意図，フレーズ構造，音の係り結び等，楽譜には直接書かれない情報や，表層構造から解析的に求められる音楽要素のこと．どちらも、先行研究に基づき、本稿内での利用に際してこのように定義する．

[7]複数の旋律が同時に演奏される音楽．現代の商業音楽は，アンビエントやニュー・エイジ等少数のジャンルの一部の楽曲を除いて，全てこれに当てはまる．

[8]2 つの旋律の音高が同じ方向に遷移すること

[9]2 つの旋律の音高が互いに逆の方向に遷移すること

[10]各パートのの全ての音高を上下に同じ量だけずらして演奏すること

[11]五線譜を 2 つ書き連ねた楽譜

[12]「カデンツ」ともいう.「いくつかの単語が結合されて『センテンス』が組み立てられるように，いくつかの和音が結合されて『和声のセンテンス』であるカデンツが組み立てられる.」[16] 第 I 巻 37 ページより

作曲，及び編曲には，和声が重要な役割を果たし，創作のどの過程においても欠かすことができないことが分かる．また，上記の 2 ステップ目以外では，作編曲活動が和声学的な要素設計のみから成り立っていることが分かる．このことから，自動作曲システムを構築するに当たり，和声の設計がいかに重要であるかが分かる．そもそも，ある程度の和声学知識を兼ね備えたシステムでなければ，作曲という行為に適切に当たることができないのである．

自動作曲システムに和声学を実装するにあたり，具体的にどの和声学を利用するかという問題がある．というのは，クラシック等における和声は，大まかな構成は世界各地で共通の認識として持たれているものの，その細かい記述の方法や，稀に見る特定の和声現象をルールとして記述するのか例外として扱うのかどうか，等といったように，流儀によって詳細は異なっているのである．本研究では，日本の多くの作曲者が基礎教育の中で共通して学習していることが知られている [17]，島岡らの『芸大和声』[13][16] を実装することとした．

芸大和声を実装した他の研究として，三浦らの *BDS (Basse Donnée System)*[23]，*SDS (Soprano Donnée System)*[24]，*MAESTRO*[25] 等が挙げられる．これらシステムには，芸大和声に記載されている「禁則」（あってはならない和音の組み合わせや和声音の並びをまとめた規則）が実装されており，入力の楽曲に適用し，それを破る箇所がないかどうかを探す形式で実行される．これらシステムは，芸大和声に記載されているようなバス課題やソプラノ課題 [3] に対して，優秀な音大生と同じレベルの性能を有していることが示されている [26]．

2.2.3 自動和声付与システムの構築

先行研究を踏まえ，本研究で構築する自動作曲システム内にも，芸大和声の諸規則を実装することとした．第 3 章で紹介する自動伴奏生成システム *ACC* の構築に先立ち，著者は卒業研究の一環として，バス課題実施 [14] システム *"Composing Music for You"*（以下 *CMY*）を実装した．図 2.2 に，バス課題の例を示す．

当初，*CMY* は芸大和声の「禁則」以外に,「標準配置の法則」や「推奨則」も実装していた [49], [50], [51]．これは，楽曲の長さ（厳密には音数）が長くなるにつれ，解候補が指数的に増加していくといった一般的な作曲の傾向 [15] に対処するために，あらかじめ探索空間を短縮するために行っていた．実際，芸大和声実装範囲のバス課題については，先行研究よりも効率よく模範解答を出力することができていた．この詳細については，著者の学士課程卒業論文に研究成果が記載されている．

しかしながら,「標準配置の法則」と「推奨則」とは和声学の初期の段階における学習の道しるべに過ぎず，学生のために便宜的に定義されたものであり，一般的な作曲には必ずしも適用されない，といった問題があった．そこで，その後は *CMY* のアルゴリズム実装形式や内部での解表現を改め，グラフ構造を利用した探索アルゴリズムを実装することにより，その実行時間を大幅に短縮することに成功した [52]，[53]．そのため，この 2 種類

[13]『芸大和声』は，著書の正式名称ではない．東京芸術大学で教えられていた和声学，という由来から，広く一般的に通称としてこのように呼ばれているため，本稿でもそう呼称することとした．
[14]バス課題等の和声課題に回答し，4 パートが全て揃った楽譜を作ることを，専門的に「実施」と言う
[15]詳細については第 2.1.2 節参照

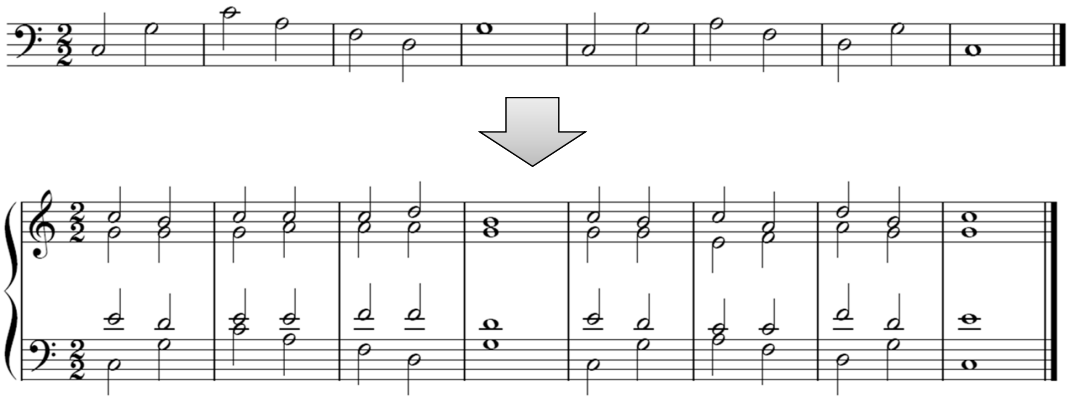

図 2.2: バス課題とその解答の例（『芸大和声』第 I 巻 p.41 課題 9–1）

の規則はその後のシステム実装には必要無くなった．以下では，卒業研究以降のこれら遷移を経た現在の *CMY* の実装と，システムの動作結果についてまとめる．

実装システムの構成

CMY は，ルールベースを利用するアプローチで，入力の旋律に自動的に和声付けを行うシステムとして開発した．和声付けとは，つまり，旋律に合った適切な和音（音程）の列を選び，さらに各和音で演奏する具体的な音（音高）を選ぶことに該当する．図 2.3 に，和声付けの具体例を示す．*CMY* は与えられた入力に対して，図 2.3 の項目 2 と項目 3 をともに自動で行う．

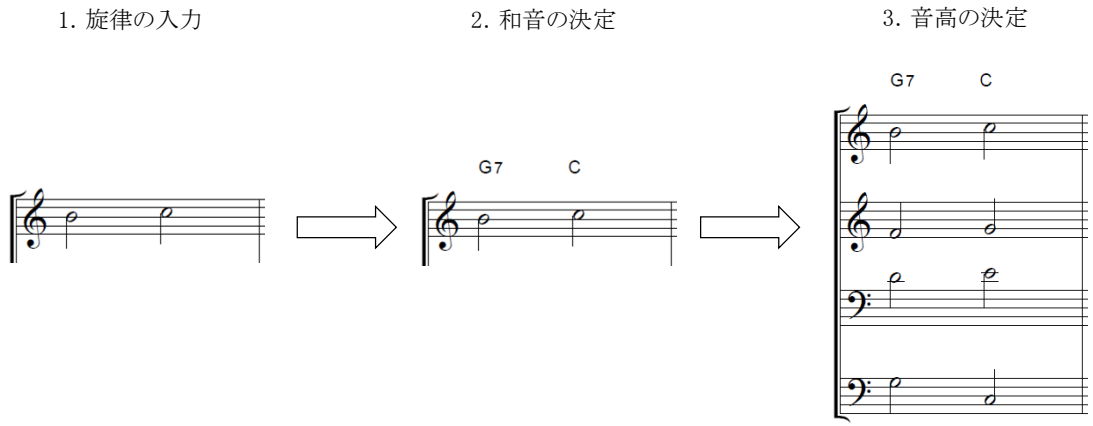

図 2.3: 和声付けの流れの例

本研究において，ルールベースを利用する手法でシステムを実装した理由は，主に 3 つある．1 つ目には，この手法は既に他の先行研究のシステムにおいて良い成果を上げているためである．2 つ目には,『芸大和声』[16] の諸規則が,「〜〜はよくない」等といったルー

ルの形で記述され，コンピュータに実装する際，ルールベースを構築する手法が最も端的であると考えたからである．そして 3 つ目に，各規則を自動作曲のアルゴリズムそのものの中に入れず，ルールベースとして別に実装することにより，後からルールの追加や変更が容易にできると考えたからである.『芸大和声』にある和声学は，あくまで古典的クラシックの和声概念をまとめたものに過ぎない．現代の商用音楽の大半に利用される和声理論がここから発展しているとは言うものの，その詳細は異なる部分もあり，今後システムをより一般的なアプリケーションとして開発していくときに，ルール内容を容易に変更できる体系を備えておく必要があると考えた．

CMY は，与えられた楽譜に対する和声付けの結果から，諸規則に反しない許容解を全探索し，その一つ以上を出力する．実装規則として,『芸大和声』第 1 巻に，学習内容の総括として記載されている「原則的な公理」から，代表的な禁則を 11 規則選び，ルールベースの形で実装した．もともと *CMY* はバスパートの楽譜を入力して四部合唱譜を作るシステムであったが，現在はバスに限らず，ソプラノ，アルト，テナーのどのパート譜を入力として与えても四部合唱譜が生成できるようになった．

システムは，入力された楽譜の各音に対して設定可能な和音を列挙し，そのそれぞれに対して他の 3 パートを定めた和音をノードとしてそろえる．その中から，和声の諸規則に反するノードを除き，各ノードから隣接する音の各ノードへリンクを張る．その後，連結された各音について，禁則の適用を確認し，違反するものはリンクを除去する．入力リンクか出力リンクのどちらかが完全になくなった和音ノードは削除され，そこから派生していた残りのリンクも削除される．その後、削除されたノードの前後のノードのリンク数が再度確認され，入力リンクと出力リンクとを共に持たないノードが全て削除されるまで確認を続ける．結果として，システムは和声音の連結可能性を有効リンクとして持った，木構造のグラフを生成する．根ノードから始まり，葉ノードで終わる経路は全て，入力の旋律に対する許容解を表し，どれも和声の禁則に反しないことになる．*CMY* が出力を決定する以上の流れを，図 2.4 にまとめた．

CMY は Java 言語で構築し，2,300 行ほどのコーディングを行った．入出力の楽譜はすべて，オープンフォーマットで汎用性の高い MusicXML 形式 [16] とした．MusicXML データ形式は，異なるプログラム間で楽譜情報を共有するために開発され，MIDI データ形式と並んで音楽情報の世界で頻繁に用いられるデータ形式である．*finale*[22] や *Sibelius*[17] 等のメジャーな記譜ソフトをはじめ，音楽の深層構造 [6] を扱った他の研究（[27] や [28] 等）でも用いられ，商業的にも研究的にも汎用性が高い．このため，ユーザは，入力の楽譜を *CMY* 特有のファイル形式で作成する必要がなく，商販されている記譜ソフトウェアで容易に作成可能である．また，出力ファイルを商販ソフトで編集することも容易である．

システム評価と考察

システムの評価は,『芸大和声』に掲載されているバス課題を入力し，出力される解の中に『芸大和声』の模範解答が含まれていることを確認して行った．入力に用いたバス課題は，教科書本文中の課題と第 1 巻巻末の補充課題から，現在の禁則実装範囲に該当する課

[16] musicXML® by MakeMusic, `http://www.musicxml.com/`
[17] Sibelius® by Avid Technology, `http://www.sibelius.jp/`

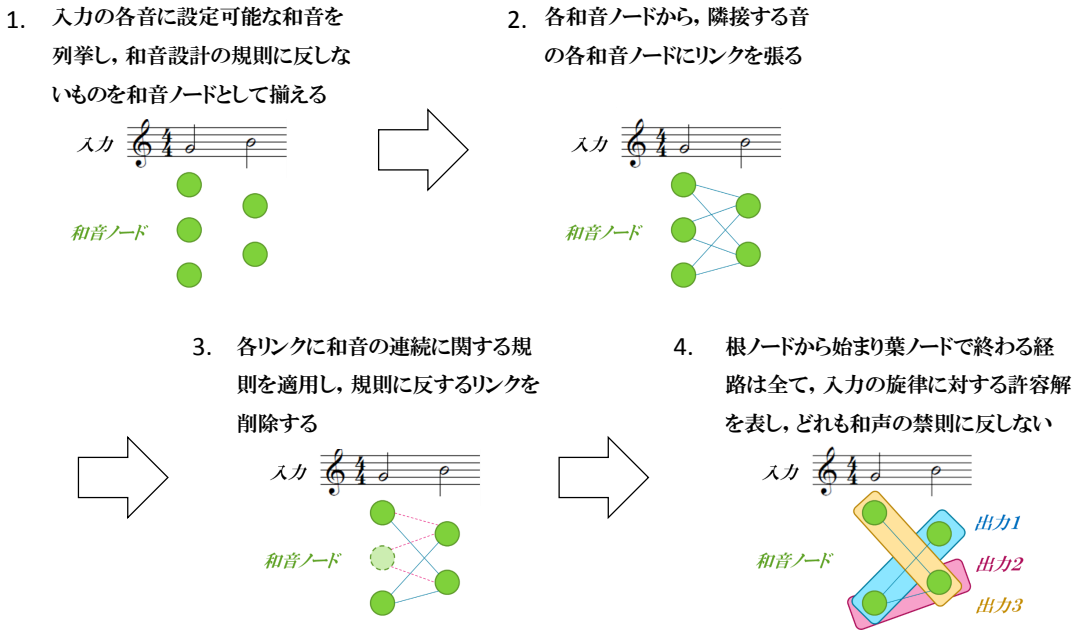

図 2.4: 自動和声付与システム *CMY* の許容解決定の流れ

題 40 題を用いた．バスパート譜以外の入力については,『芸大和声』掲載のバス課題模範解答から，バス以外のパートを抽出し，入力として与えた．その際，生成される解の中に教科書の模範解答（入力の楽譜の元となったバス課題解答）が含まれていることを確認した．

CMY は，3 和音の第 1 転回系までを扱える程度に実装している．禁則は 11 規則を実装しているが，教科書の未実装部分に登場する第 2 転回形や 4 和音を用いる課題が入力されると解を探索できない．現在扱える課題については，許容解が多いもの（数百程度）でも，数秒以内に探索を終えることができている．

CMY は性質上，厳格に和声学に従うような自動作曲課題を解く専門的なシステムとして構築されている．これは音楽大学の授業における教育支援システムや，課題自動採点等への応用に適しているであろう．木構造を利用した内部表現により,『芸大和声』記載の規則を全て実装した場合も，短時間での解生成を維持できることが期待される．

しかし，許容解の中から優れた解，すなわち教科書に掲載されているような「模範解」を選ぶ機構は *CMY* には存在しない．それは許容される解の優劣の決定についての言及が，和声学には乏しいからである．和声の規則は，存在してはならない和音や和音の連結については詳細に規定している．しかし，最終的な解決定に必要な，許容解同士の優劣の評価は，実際に教師の提示する楽譜や音源等を元に，学習者が、それぞれ各自で形式化して覚えていく学習形態になっている．許容解に対して付けられる各作曲者個人の優劣判断基準が，その作曲者の個性を表し，芸術における多様性を生み出していると考えられているのである．

つまり，*CMY* は和声学の学習システムとしての応用は期待されるものの，より普遍的な自動作曲システムへの利用には不向きである．和声学を実際のアプリケーションに応用する際には，その規則を作曲の過程そのものに利用する他に，厳格さを求めないより柔軟

な方法で記述し，許容解の中での優劣評価に利用できればなお良い．和声学の自動作曲への応用について，著者は第 3 章で別な実装法を提案し，評価と考察を行っている．詳細については，その項を参照されたい．

2.3 印象誘導としての作曲 - 生体心理学の研究

我々は楽曲を聴くと,「この曲は好きだ」や「この曲はもう聴きたくない」等，楽曲に対する印象を形成する．この印象形成は意識して行うこともあるが，無意識的に行われる場合も少なくない．特に商用音楽は，顧客により多く購入されるためにと，消費者である聴衆に好印象を持ってもらうことを意図して楽曲を創作していると考えられる．また，自己表現としての個人的な作曲活動も，作曲者自身が持っている「なにか」を表現しようと，作曲者が意図的に楽曲を作っていると考えるのはおかしくない．

このように，楽曲の作曲者は必ず何かしらの「意図」を持っている．音楽の様々な要素を組み合わせて一つの楽曲を作り上げる創作行為は，その「意図」を表現する手段として行われ，その「意図」は完成された楽曲作品の中に存在し，集結する．後から楽曲を聴くリスナーは，楽曲を通して表現されている作曲者の「意図」を意識的に解釈したり，または無意識に理解したりして，楽曲から何かしらの「印象」を抱く．

自動作曲研究では，リスナーの耳に伝わる音楽の物理的現象や，記譜した際の他の楽曲との類似点，または音楽学的な楽曲構造の仕組みを解析して曲を作り上げる研究が多い．しかし，音楽鑑賞に付随する，上述したような作曲者意図やリスナー印象の形成に着目し，それを作曲機構に取り入れたものは少ない．本節では，その作曲者意図とリスナー印象との関係に着目し，そもそも自動作曲システムに実装可能かどうか，また，もしそうであれば，どのようにして計算機上に実装できるかを検討する．

2.3.1 先行研究

音楽によって誘発される情動に関する研究は，すでに 100 年以上も前から行われている [29]．これらは，被験者が楽曲を聴いて感じたものをまとめる研究（[30]，[31]，[32] 等）と，被験者が楽曲を聴いて，その楽曲が伝えようとしていると感じたものものをまとめる研究との 2 つに大別される [33] が，この 2 つをあえて区別しない研究も少なくない（[34]，[35]，[36] 等）．音楽情動の研究において，リスナーが感じる印象と作曲者が意図した印象とは，一般にはっきりと区別されていない．

生体心理学観点からの音楽情動研究の多くは，研究対象として「楽曲」と「楽曲リスナー」の 2 者を対象として取り上げ，その関係性について考察する（[37] 等）．全ての楽曲の裏には，それを創作した「作曲者」が必ず存在するのであるが，この研究分野において作曲者を定義し，楽曲を媒体とした作曲者とリスナーとの関係性に着目した研究は極めて少ない．研究内で扱っている音楽のジャンルや楽曲の形態は本稿で扱っているものとは異なるが，作曲者を扱った数少ない先行研究の例を以下に紹介する．

Baraldi ら [38] は，MIDI ピアノ単音（鍵盤一つ）の弾き方を変えることにより，演奏者（作曲者）が様々な感情をどのように伝えるか，またリスナーがそこから何を感じているかを調べている．また，Weale[39] は，実環境で収録した音やシンセサイザで生成した音

等を混ぜた音楽ジャンル EM（Electroacoustic Music）の作曲者意図とリスナーの印象評価を比較する検証方法を提案している．EM はそもそも新しいジャンルの音楽であり，それを聴いたことのある一般聴衆は少ない．Weale は，楽曲の題名や作曲者が作曲する裏で考えていたストーリー等を徐々にリスナーに開示し，それによって，リスナーが楽曲を聴いて抱く印象がどのように変化していくかを観察した．これらのように，作曲者意図を取り上げる研究は限られており，その対象とする音楽は他の音楽情動研究や自動作曲研究で扱われる，一般に聴衆の多いジャンル（クラシックやポップス，ジャズ等）ではない．特定の音楽ジャンルにおいて，極限られた楽曲しか扱われていない．

似たような研究の例として，演奏者の意図について行われたものを挙げることができる．De Poli ら [40] は，あるバイオリンの演奏を一音一音解析し，それぞれの音が持つ音楽的特徴の違いと，演奏者が伝えようとした意図との関係を解析している．しかし，その実験系には演奏者以外の被験者はおらず，実際にリスナーにどのように印象が伝承したかに関する検証は行われていない．

音楽情動分野において作曲者意図を扱った研究が少ない原因として，いくつかの理由が考えられる．一つに，研究で用いられる楽曲が過去に書かれたものである場合が多く，作曲家にその意図を確認する術がない等といった，作曲者の意図を調査することが難しい実態が挙げられる．特にクラシック音楽等は，著作権等の面から研究で使いやすいという利便性を持つ反面，作曲者や作曲当時の社会環境について調べることが困難であるという短所がある．近年は学術用の音楽データベース [18][41] 等も構築されているが，これらはあくまで楽曲の表層構造 [6] のみを集めたデータベースであり，その楽曲に込められた作曲者の意図に関する情報を明示的に保持しているものではない．

2.3.2 作曲による印象誘導実験

生体心理学観点の研究において，楽曲とリスナーの関係性に着目した先行研究が大多数を占めているわけだが，著者は自動作曲システムを構築することを視野においているため，作曲者も定義した予備実験を行うことにした．具体的な興味として，作曲者は楽曲を介して，どの程度リスナーの心理活動に影響を及ぼしているのかを探った．

楽曲がリスナーに心理学的な影響を及ぼすことがある，ということは，前述した先行研究から分かっている．また，それらの心理学的反応を引き起こすと考えられる楽曲の特徴をまとめた研究もある [42]．これらの研究は，様々な現象をまとめる際に，リスナー視点から楽曲を解析している．つまり,「リスナーはこの楽曲群を聞いてこれこれと反応しているが，では，その楽曲に共通するものは何なのだろう？」というように，結果から因果関係を構築している．しかし，それとは対照的な，そこで導いた因果関係を再現できるかどうか，作曲者視点から考察した研究，つまり,「リスナーにこれこれの反応をさせるために，こういう楽曲を構築しよう」という誘導的な研究は見られない．

実際の作曲者は，聴衆の属性や聴衆の嗜好，楽曲が演奏される場所やその演奏形態等をある程度見据え，それに応じて作曲をしている [19]．作曲者が作曲をする際に設定する制約

[18] RWC 研究用音楽データベース　`http://staff.aist.go.jp/m.goto/RWC-MDB/index-j.html`

[19] 著者が教育大学の教授（プロの作曲家），クラシックの声楽アーティスト，作曲専攻の大学生等と本研究について相談する中で得た見解．作曲者が依頼されて行う作曲は,「どこどこ主催のイベントのため」や「これこれに関する試験のため」,「これこれのドラマのこんな場面のため」等と，多くの制約が与えられて作曲

や，楽曲を通してリスナーに伝えたいメッセージ等，楽曲制作の裏に存在する漠然とした目的を，本研究では「意図」と呼ぶ．また，作曲者は必ず意図をもって作曲していると考える．その作曲者の意図を理解し，システム内に実装することができれば，作曲者が行うようなリスナーの印象誘導をシステムに模倣させることが可能となることが期待される．

本節では，著者が研究開始前に自作し，既に複数回演奏していた楽曲を用いて実験を行った．まず，楽曲を聴くリスナーに感じてもらいたいと著者が事前に設定していた感情（意図）と，実際に楽曲を聴いたリスナーの印象とを比較した．その後，特定の感情を更に誘起するために，先行研究に基づいて楽曲を編曲し，音楽的特徴を変更し，再度リスナーたちの印象を調べた．限られた例ではあるが，作曲者は楽曲の音楽学的特徴を意図的に変更することにより，従来研究でまとめられていた楽曲によるリスナーの印象形成を機械的に誘導できることを確認した．この実験結果を踏まえた具体的な展望として，この知見をアルゴリズム的にシステム内に実装することにより，作曲者によるリスナーの印象誘導を模倣した自動作曲システムを構築したいと考えている．

実験概要

本実験では，被験者に 2 日にわたって楽曲（著者が実験前に作曲・演奏していた自作の楽曲）を聴いてもらい，楽曲に対するリスナーの印象をアンケートをもって評価した．実験後，著者はアンケートの結果を踏まえて楽曲を編曲し，編曲したその楽曲を先の実験と同じ被験者に再び聴いてもらい，リスナー印象の変化を再度アンケートを利用して評価した．

なお，楽曲を聴いている間は脈波，2 種の皮膚電気活動，及び指尖脈波も計測した．結果として，個人差は大きいものの，楽曲を聴いている際のリスナーの生体反応には統一的な傾向が見受けられたことや，その一部は作曲者（著者）自身の生体反応にも観察されたこと等が挙げられる．楽曲の開始・終了，また旋律の変わり目で，作曲者もリスナーも反応を示す場合が多く，また楽曲の特徴に応じて異なる傾向の反応を示していることも観察した．ただし，紙面の関係上，これら生体反応の解析結果は本稿では扱うことができない。その詳細については [45], [46]，及び [47] を参照していただきたい．

また，本実験では毎回，被験者にそれぞれの楽曲を最初から最後まで全て聴いてもらった．先行研究の多くは，実験時間短縮及び複数楽曲に及ぶ実験環境統制のため，被験者に楽曲の一部のみ（冒頭 30 秒等）を聴かせたうえで印象を評価している（[35] 等）．本研究では，作曲者は意図をもって作曲する際，楽曲全体を一つの作品として作り上げ，その全体に意図を込めていると考えた．特に，ある楽曲の全体を聴いたリスナーと，その一部だけを聴いたリスナーとでは，その抱く印象が異なりうることは優に考えられる．そこで本実験では，全ての被験者に楽曲の全体を聴いてもらい，楽曲に対して抱いた総合的な印象をその楽曲に対する評価として扱うことにした．

を行う．趣味で作曲を行う場合も，自己表現のほかに,「映像の BGM 用に」や,「最近聞いた新たな曲調を取り入れてみよう」等と，何かしらの目的をもって作曲する場合が大半である．

実験設定

被験者（32 人，男性 22 人，女性 10 人，平均年齢 23 歳）には，2 日にわたり，それぞれ同じ実験環境で，実験者の書いた楽曲を 2 つずつ，合計 4 曲を聴いてもらった．各楽曲の開始前には鳥のさえずり等が聞こえる環境音 [20] を 1 分程度聴かせ，被験者が安静した状態で楽曲聴取に臨めるようにした．なお，各日に流す最初の 1 曲は被験者が実験環境に慣れるため等に利用し，印象誘導に関する実際の解析は全て 2 曲目として流した楽曲 2 曲にのみ行った．全ての音源を聴き終えた後，被験者には楽曲に対する印象等を問うアンケートに回答してもらい，2 日目の実験終了後には，楽曲や実験に関する総合的な印象をインタビュー形式で尋ねた．実験は 1 回につき，20 分程度かかった．

被験者アンケートには，曲の印象を書く自由記述欄と，9 つの感情因子について「○○である」から「○○でない」までを，5 段階で答えるリッカート尺度を用いた設問を設けた．音楽情動の印象評価実験では様々な感情モデルが利用されているが，本研究では Bruner が，主に 7 つの先行研究を元に選定した感情因子の集合 [42] を利用することにした．9 つの感情因子と本研究で利用した訳語，及び Bruner がまとめた，各感情を誘起する音楽的要素とその特徴を表 2.2 に示す．

表 2.2: Bruner の 9 つの感情因子 [42] と，本研究で用いた訳語，及び各感情を誘起する音楽要素とそれらの特徴

	Serious	Sad	Sentimental	Serene	Humorous	Happy	Exciting	Majestic	Frightening
	重々しい	悲しい	感情的である	穏やかである	ユーモアがある	嬉しい	興奮させられる	威厳がある	恐ろしい
Mode	Major	Minor	Minor	Major	Major	Major	Major	Major	Minor
Tempo	Slow	Slow	Slow	Slow	Fast	Fast	Fast	Medium	Slow
Pitch	Low	Low	Medium	Medium	High	High	Medium	Medium	Low
Rhythm	Firm	Firm	Flowing	Flowing	Flowing	Flowing	Uneven	Firm	Uneven
Harmony	Consonant	Dissonant	Consonant	Consonant	Consonant	Consonant	Dissonant	Dissonant	Dissonant
Volume	Medium	Soft	Soft	Soft	Medium	Medium	Loud	Loud	Varied

実験両日に 2 曲目として流した楽曲 2 曲を，それぞれ楽曲 A，楽曲 B と呼ぶ．これらは，どちらも著者が研究開始前に作曲し，2012 年 8 月にアマチュア・グループが制作・公演したミュージカルにおいて演奏された楽曲である．本実験においてミュージカルの演奏曲を利用した理由は，作曲者の意図が他の様々な音楽に比べてより明確に，また客観的に表現されていると考えたためである．ミュージカルで演奏される楽曲は，特定の場面や情景を具体的にイメージして作曲される．そのため，その楽曲に込められた意図や，作曲者が楽曲を通してリスナーに求めた印象等は，言語的に表現することが他の楽曲と比べて容易であると考えられる．

楽曲 A，楽曲 B は，それぞれ 2 分半程度に編曲し，MIDI 音源で演奏したものを mp3 形式のファイルとして利用した．楽曲はともに *Finale2010J*[21] で作成され，*Garritan Instruments for Finale*[22] を用いて WAV ファイルに出力された後に，mp3 形式に変換された．著者は被験者実験に先立ち，それらの楽曲を作成する際に聴衆に感じてもらいたいと意図していたリスナー印象を，被験者アンケートと同じ Bruner の 9 つの感情因子を用いて，5 段階

[20] YouTube, Forest and Nature Sound 10 Hours, `http://www.youtube.com/watch?v=0dIJ2x3nxzQ`
[21] finale® by MakeMusic, `http://www.finalemusic.com/`
[22] Garritan® by MakeMusic, `http://www.garritan.com/`

で評価した．

楽曲 A は，音楽学でいう「aba」の構造をしていて，ハ短調，8 分の 12 拍子であった．Garritan の MIDI 音源「Steinway Piano」で収録され，中間部「b」では，一部ホ長調に変調した．楽曲 B は「ababa」の構造をしていて，変ロ長調，8 分の 12 拍子であった．MIDI 音源「Choir Ahs」が利用された．2 つの楽曲の MIDI 音源には，それぞれがミュージカルの中で演奏された際の演奏形態（それぞれピアノ独奏，アカペラ合唱）に最も似ていると思われる音源を利用した．被験者の半数は 1 日目に楽曲 A を聴き，半数は 2 日目に楽曲 A を聴いた．

2 日間の実験終了後，著者は楽曲 A について得られたアンケート結果を踏まえ，作曲者意図とリスナー印象の差が特に大きかった感情因子の評価を，当初の意図により近づけるために，楽曲 A を再編曲した．再編曲は，表 2.2 にある楽曲特徴を調整することにとどめ，曲調の大きな変化等は行わなかった．この追実験は最初の実験の凡そ 1 か月後に実施し，最初の実験と同じ手順で被験者の印象を評価した．被験者は 12 人で，全員最初の実験にも参加していた．なお，便宜のため，この再編曲した楽曲は楽曲 C と呼ぶことにする．

実験結果

図 2.5 に，各楽曲に対する作曲者の評価（意図）と，被験者の評価（リスナー印象）の平均をプロットしたものを示す．作曲者は，楽曲 A に対して凹型，楽曲 B に対して凸型の評価を意図していたが，リスナーもそれに則した印象を形成していることが分かる．

また，楽曲 A と楽曲 B に対する被験者の印象のみをプロットしたものを図 2.6 に示す．2 つの楽曲に対する被験者印象の分散分析（1 要因被験者内）の結果から，重々しい（$F(1,31) = 90.49, p < 0.01$），悲しい（$F(1,31) = 53.53, p < 0.01$），感情的である（$F(1,31) = 11.52, p < 0.01$），穏やかである（$F(1,31) = 19.04, p < 0.01$），ユーモアがある（$F(1,31) = 38.16, p < 0.01$），嬉しい（$F(1,31) = 54.51, p < 0.01$），興奮させられる（$F(1,31) = 12.50, p < 0.01$），威厳がある（$F(1,31) = 67.85, p < 0.01$），恐ろしい（$F(1,31) = 87.28, p < 0.01$）の全てについて，それぞれ強い有意差があることが示された．すなわち，各感情因子に着目した時に，被験者はこの 2 つの楽曲に対して，全く異なる印象を抱いていたことが分かる．

最後に，楽曲 C の評価に携わった被験者の印象を，楽曲 A と楽曲 C について，楽曲 A に対する作曲者意図とともにプロットしたものを図 2.7 に示す．実験者は，作曲者意図とリスナー印象が特に異なっていた感情因子のうち,「嬉しい」と「恐ろしい」に着目し，これらのリスナー印象を作曲者意図により近づけようと楽曲 A を編曲し，楽曲 C を作成した．具体的には「嬉しい」の印象を弱め,「恐ろしい」の印象を強めるために，表 2.2 を参照し，Pitch, Rhythm と Volume を変更した．

Bruner によると，Pitch（楽曲内各パートの音高）は「嬉しい」では高く,「恐ろしい」では低くなる傾向がある．楽曲 C では，高音を演奏する旋律のオクターブを下げる等して，音程は変更せずに一部の音高だけを低くした．Rhythm（旋律のリズムや曲の拍）に関しては,「恐ろしい」曲では不均等である傾向がある．楽曲 C では小節ごとにテンポを変化させたり，旋律に突如シンコペーション [23] を加えたりする等して，不均等なリズムを

[23] ある旋律が，小節の強拍・弱拍構造を無視して演奏される技法

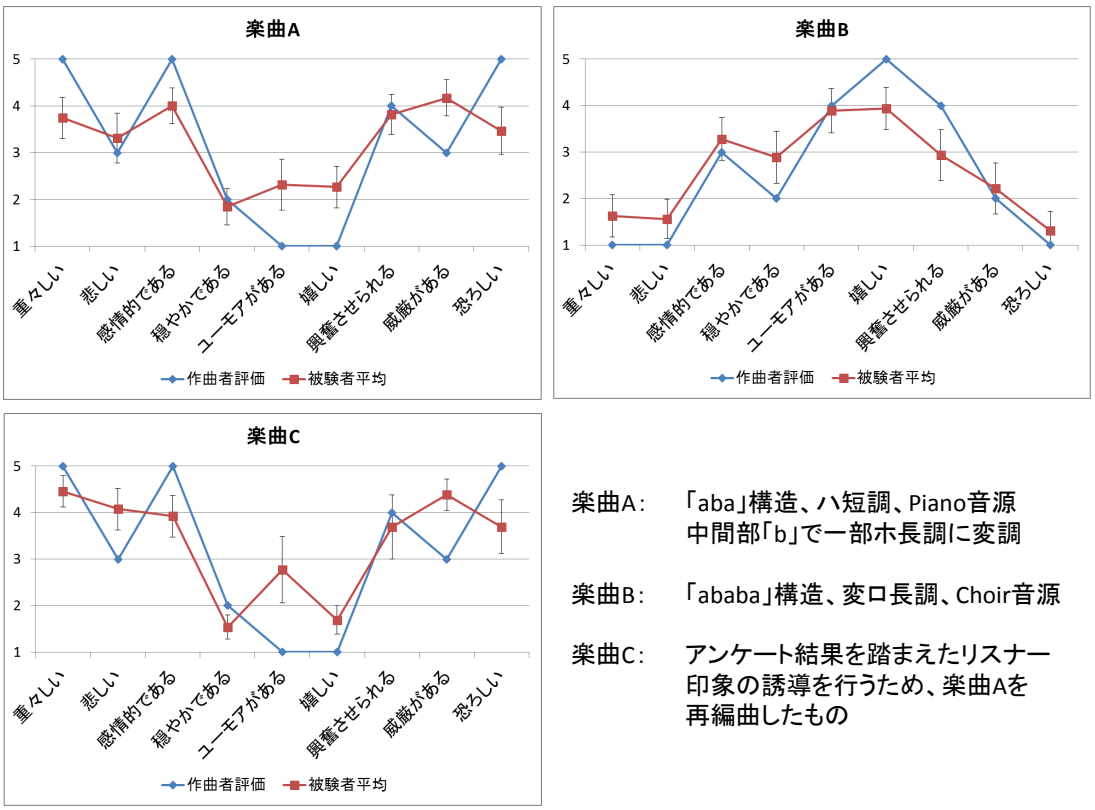

図 2.5: 楽曲 A,B,C に対する作曲者の評価と，被験者（32 人）の評価平均（エラーバーは標準偏差を表す）

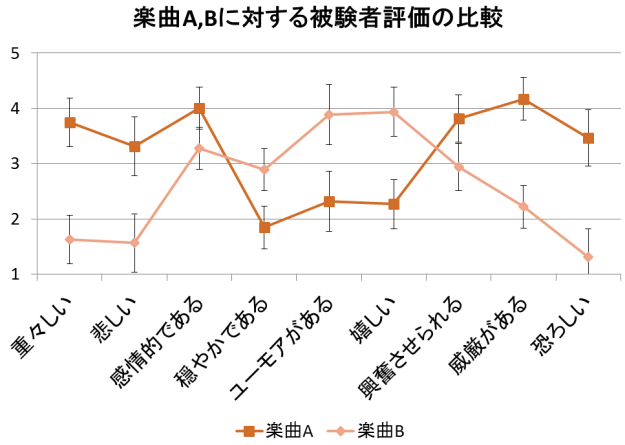

図 2.6: 楽曲 A,B に対する被験者（32 人）の評価平均（エラーバーは標準偏差を表す）

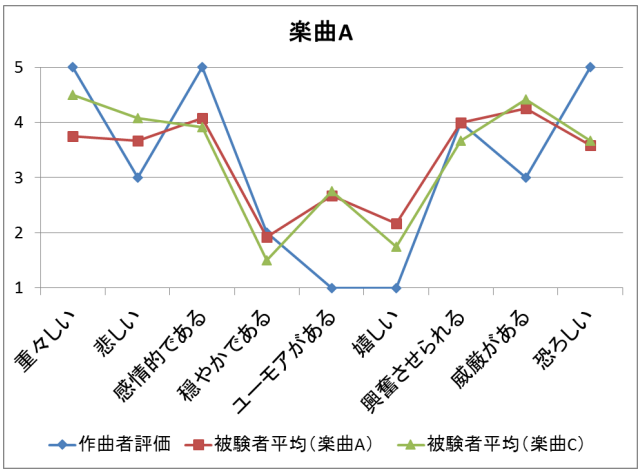

図 2.7: 楽曲 A に対する作曲者評価と，楽曲 A,C に対する被験者（12 人）の評価平均

増やした．Volume（音量）に関しては，強弱をところどころ変化させ，楽曲全体を通して音量が不均等になるように編曲した．編曲後の楽曲 C は，小節数が変わっていないものの，リズムやテンポに変化が生じたため，編曲元の楽曲 A よりも 30 秒程長くなった．

楽曲 A と楽曲 C に対する被験者 12 人の印象平均の分散分析（1 要因被験者内）の結果から，重々しい（$F(1, 11) = 3.86, p < 0.1$）についてのみ，有意傾向があることが示された．編曲の狙いだった「嬉しい」や「恐ろしい」の印象に関しては，リスナーの中での印象の変化に有意差は示されなかったが，傾向として，被験者平均が作曲者の意図に近づいたことが伺える．全体として，被験者は負の印象を強め，正の印象を弱めた評価をしている傾向があった．

考察

図 2.6 にあるように，楽曲 A と楽曲 B とに対するリスナーの印象は，有意に異なっていた．また，図 2.5 から，それらリスナーの印象が，作曲者の意図していた印象に類似していることも明らかとなった．被験者群や実験に利用した楽曲がまだ限られていることから，すべての音楽に共通する一般的な結論が導き出せるわけではない．しかし，少なくとも一部の音楽において，作曲者は実に，リスナーの印象形成を誘導することができていると分かる．特に，ミュージカルといった歌劇風芸術では，この印象誘導を効果的に利用し，観客にメッセージを伝え，臨場感を生み出しているだろうことが推測される．

また，図 2.7 から，作曲者は楽曲の音楽的要素を変更することにより，同じ原曲を用いながらも，リスナーに異なった印象を誘起することが可能であることが見える．今回の実験の編曲で，実験者は「嬉しい」と「恐ろしい」の 2 つの感情因子についてリスナーの印象を誘導しようとしたが，結果として「重々しい」の軸で有意傾向のある結果が得られた．これは，楽曲から感じる印象のラベル付けに匹敵する聴衆の心理活動が，各個人の間で差を持つことに起因すると考えられる．アンケートを通して具体的に説明を求められた時に，

リスナーは「嬉しい」と「恐ろしい」においては，編曲の意図に従った印象の変化を示していなかったものの,「重々しい」における印象の変化は，作曲者の意図に近いものであった．つまり，細かい印象誘導は難しいとしても，楽曲全体に関する総合的なリスナーの印象を，作曲者による具体的な楽曲特徴の変化で誘導することは可能なのである．

従来の自動作曲システムは，作曲した楽曲に対するリスナーの印象をフィードバックとしてシステムに与え，それをもとに内部の作曲パラメータを変更したり，その楽曲に具体的な変更をもたらしたりする [15]．楽曲に対する作曲者の意図は考慮されず，リスナーが楽曲に対して抱く印象も，一度作曲を終えてから事後的に検討される場合が多い．しかし，リスナーが楽曲を聴いた際に抱く印象は，リスナーに楽曲を聴かせる前に，事前に作曲の段階で，音楽要素を具体的に設計することにより，ある程度誘導できることが分かった．従来の自動作曲システムでは，この印象評価部分が完全にユーザの主観的事後評価に任されていたが，先行研究でまとめられている，音楽学などの一般的な知見を作曲機構内に予め実装することにより，ユーザが評価に充てる時間や作業量を減らし，より簡単にその狙いの楽曲を生成することが可能であると考えられる．

2.4 自動作曲研究の課題の総括

第 2.1.1 節でみたように，自動作曲の課題に対して，様々なアプローチからの研究がなされている．それぞれのアルゴリズムには，その長所・短所が存在するが，システムが生成する楽曲の質は，少なくともシステムに実装されている音楽学の知見に大きく影響されるということが分かっている．特に，本稿で扱っているクラシック等の音楽ジャンルにおいては，大多数の聴衆に普遍的に理解され，共有されている，楽曲の要素や構造，またそれらから派生する印象形成過程などが存在するため，それらに関わる知識をいかにシステムに実装するかが大事となる．特に，将来的なシステムの拡張性を見据えて，音楽学の知識がアルゴリズムそのものに内蔵されない実装が必要となる．それと同時に，既存の楽曲の真似に留まるような作曲ではなく，まったく新しい楽曲を自動的に生成可能である柔軟さがシステムには求められる．

また第 2.1.2 節で議論したように，自動作曲の最大の難点は，その膨大な解空間から，ユーザ [24] が意図した楽曲を求めるところにあると考えられる．特に実際のアプリケーションへの応用を考えたとき，作曲システムが実時間内に実行できなければいけない場合が多い．しかし，著者が調べた限り，システムの実行時間を議論し，特別に考慮した先行研究はない．楽曲の長さに合わせて実行時間が指数的に伸びていく自動作曲の課題では，ユーザの入力に応じて作曲システムが効率良く探索を行わなければいけない．

そして最後に，ある一つのシステムで多様な楽曲を生成可能であることが望まれる．特定の作曲課題に特化したシステムを作るのは学術的には興味深いが，実際にユーザに利用される状況を考慮すると，実用化に乏しいという考えもある．生成する楽曲の一部限定は避けられないとして，その設定した範囲内で，音楽的に面白く [25]，特徴が有意に異なるよ

[24] それがシステム設計者になるか，システムを利用する一般の聴衆になるか，その他の第三者になるかは研究によって異なる

[25] 音楽的面白さというものの定義は曖昧で，解釈する者によってその意味が大きく異なる．ここでは，システムを利用するユーザが飽きないように，それぞれ特色のある様々な楽曲を生成できる，という意味で「面白い」と述べている．

うな出力が生成できなければいけない．

以上をまとめると，自動作曲システムとその研究に求められる点として，以下の 3 つが挙げられる．

- 音楽学の知識が，システムに内蔵されない，容易に拡張可能な形で実装されること
- 膨大な解空間を適切に探索し，実時間内に出力を生成可能であること
- 同一のシステムから，特徴が有意に異なる様々な楽曲が生成可能であること

著者は，これら課題を同時に克服可能な実装体系を提案し、それを具体的に，自動伴奏生成システムという形で実装評価した．システムの実装と評価結果について，著者は一度 [48] にて発表を行っているが，発表後もシステムの拡張は続けてきた．以下第 3 章以降では，著者の構築したそのシステムについて詳細に説明し，自動伴奏生成に限らず，自動作曲研究一般への貢献を紹介する．

なお，自動作曲研究における他の課題として,「音楽」の性質上，ある楽曲（もしくは自動作曲システム）に対して行われる評価がどうしても主観的になる場合が多い，というものがある．先行研究の中には，様々な作曲システムに対する何か一貫した評価指標を示したものはない．音楽情報の研究においては，各研究ごとに自動作曲の意義が議論され，その評価体系も独自に検討されるため，先行の研究に基づいたシステムの構築・発展が難しいのが現状である．

以降本稿では，実装したシステムの特性に基づき，自動作曲分野の中でも，特に限定された自動伴奏生成について議論をしていく．しかし，その提案アルゴリズムによって克服しようと目標設定した課題や，それに向けた実装方針等は，以上で議論してきた，自動作曲研究一般に該当するものばかりである．本研究の目的は，一度，対象とする研究範囲を狭めることにより，提案する機構の有用性を検証するところにあり，そこで得られた知見を踏まえて，今後，より一般的な自動作曲システムを開発する予定である．

第3章　実装システム

本章ではまず，実装した自動作曲システムの概要を説明する．その後，システムの2つのモジュールをそれぞれ詳細に説明していく．その際，楽譜や数式等を用いて実装した諸規則を説明し，それぞれ，期待されるシステムの挙動や性能についてまとめる．

3.1　概要

本研究では，自動伴奏生成システム*"Auto Chorus Creator"*（以下*ACC*）を構築した．ユーザは*ACC*にリード譜を入力し，*ACC*はそれに残り3パートを追記した四部合唱譜を出力する．ユーザは，出力に求める表層的な楽曲特徴やシステム内部の評価に使われる音楽学要素を指定することもできるが，それらの設定に一切触れずとも，様々な出力をシステムから生成することも可能である．*ACC*は，そのまま演奏に利用できるスコア譜やパート譜を生成するが，ユーザが出力された楽曲を土台にして更なる編曲を加える，などといった作曲補助ツールとしての応用も想定している．つまり，*ACC*は自動伴奏生成システムであり，作曲初心者へのサポートシステムでもある．

*ACC*は元々，2014年2月に著者が参加した，独立行政法人産業技術総合研究所（AIST）[1]主催の学生実習の一環として構築されたものである．その際の研究成果について，著者は2014年9月の音楽情報処理研究会のトップカンファレンスInternational Computer Music Conference (ICMC) とSound and Music Computing Conference (SMC) の合同会議において発表をした[48][2]．本稿では，その発表で得た知見を踏まえ，さらに改良を重ねたシステム*ACC*について紹介する．

*ACC*のユーザ設定

*ACC*の対象とするユーザとして，作曲技術を持たないが，作曲はしたいと思っているユーザを設定する．このユーザ設定の裏には，最近の動画投稿サイトや音楽制作に見られる傾向を踏まえた考察がある．世界的に広く人気を呼んでいる動画投稿サイトYouTube[3,4]や，日本での登録者数が多いニコニコ動画[5]等のトレンドを見ると，最近は投稿された動画の中でも特に音楽ビデオ等に別項目を設けることにより，検索をしやすくしようとする

[1]産総研 `http://www.aist.go.jp/`

[2]Joint Conference ICMC—SMC 2014 `http://www.icmc14-smc14.net/`

[3]YouTube `http://www.youtube.com`

[4]Mushroom Network社のまとめによると，YouTubeは月ごとの検索数がGoogleに次いで世界2位であり，Yahoo!やbing等，他の有名な検索エンジンを合わせた数よりも多くのユーザを保持している．`http://www.mushroomnetworks.com/infographics/youtube---the-2nd-largest-search-engine-infographic` 2015.1.14現在

[5]niconico `http://www.nicovideo.jp/`

工夫等が見受けられる．動画投稿サイトには，プロの音楽家たちのミュージックビデオやプロモーション動画等が多数あげられていると同時に，一般ユーザによるカバーソング[6]の制作や，独自で創作したオリジナルソングの動画投稿等が著しく増えている．こうした中で，一般ユーザの音楽に関する興味は，従来の受動的な聴取行為から発展し，今は能動的な制作行為までも含むようになっている．

しかし，作曲という行為は，音楽の様々な要素（旋律，和声，拍子等）を同時に扱う，複雑な芸術創作活動であり，それを専門職として行っている人は，多くの訓練を経てその作曲を行っている．一般的なユーザも，本人が弾いたことのある限られた楽器で旋律を奏でたり，またはリズムを作り出したりすることはできても，バンドや楽隊用に，同時演奏される複数のパートを持った複雑な楽曲を作り上げることは難しい場合が多い．本人の勘やセンスで作曲を行う人も多いが，本稿著者のように，作曲に関する専門知識を習得し，より効率良く自分の求めている「音」を作り出したいと思っている人も少なくないだろう．

第 2.3 節で紹介したように，作曲者は楽曲の表層構造[7]を変更させることにより，リスナーが楽曲に対して抱く印象を誘導することが可能である．自分のためだけに作曲をするようなユーザは，自身さえ楽曲を好めばいいので，ここまで検討する必要がないかもしれない．しかし，自身が創作した楽曲を他人に発信し，何かしらの共感や感動を抱いて欲しいと考えているユーザの場合，同じような楽曲を聴いてきた者同士が，無意識のうちに構築している楽曲の解析機構や，どの音楽要素によってどの印象をどのように誘発する事が可能か等についての知識が必要になる．昨今増えてきているこのような「一般的な作曲者」は，作曲に関するこのような専門的な知識を持っていない人が多いと考えられるため，それをある程度サポートできるシステムが必要となる．

本稿で紹介する自動伴奏生成システム *ACC* は，具体的な印象を扱う段階までは到達していない．しかし，より一般的な自動作曲システムへと今後拡張していくためのプラットフォームとして，本システムは構築できている．本システムのユーザとして,「最低限の音楽知識は持っているが，専門的な作曲技術を持っていないユーザ」を設定しているのは，以上のような理由による．

システム設計指針

第 2.4 節で，自動作曲システム研究に求められる性質を 3 つ上げた．実用的なアプリケーションを見据えた自動作曲システムは，一般に

- 音楽学の知識が，システムに内蔵されない，容易に拡張可能な形で実装されること
- 膨大な解空間を適切に探索し，実時間内に出力を生成可能であること
- 同一のシステムから，特徴が有意に異なる様々な楽曲が生成可能であること

であることが求められる．

そこで，*ACC* はまず，音楽学の知識はルールベースとして実装し，必要に応じて作曲に使うルールを簡単に選択できるようにした．一点目にあるように，ルールは後から容易

[6] カバー（ソング）: ポピュラー音楽の分野で，他人が発表した曲を演奏・歌唱して発表すること.（Wikipedia より `http://ja.wikipedia.org/wiki/\%E3\%82\%AB\%E3\%83\%90\%E3\%83\%BC`）

[7] 第 2 章脚注 6 参照

に変更可能である形式で実装し，アルゴリズム自体に音楽学の知識を内在させないようにする．具体的には，第 2.2.3 節で説明した，著者が構築してきた和声課題実施システムにも実装した，和声学の諸規則をルールベースとして再実装し，システム内の探索で求た解の評価に利用することとした．その際，和声学の諸規則は連続関数として実装し，従来の実装手法では同等に優れている（劣っている）と評価されたような二つの解に対しても、より客観的に優劣の差を付けれるように設定した．

二点目の実時間実行については，システム内の探索機構に制約を設けられるようにする，という解決策が考えられる．その際，ある一定値で探索を打ち切った場合にも，それなりに評価の良い解を求められているアルゴリズムを実装する必要がある．本研究では，他の作曲研究でも用いられているダイナミック・プログラミングの手法 [43] を取り入れ，必要に応じて適度に探索を打ち切れるようにした．

最後に，三点目の異なる楽曲の生成について検討する．複数の専門家にある旋律を渡し，「この旋律に合う伴奏を作曲せよ」といった場合，専門家たちはそれぞれ異なった伴奏を作曲するであろう．同じ旋律といった共通の制約があることから，彼らの伴奏には似たようなコード進行が用いられたりする可能性はある．しかし具体的な伴奏音一つ一つの音高やリズムまで比較した場合，彼らの作曲する伴奏は，多くの点で異なったものになるであろうことが推測される．

一般に，ある旋律に対して，同等に優れている伴奏（解）は無数に存在すると考えられる．これを数学的に捉えるなら，作曲の解空間における評価関数は，一つの最適解に向かって単調増加していくような形ではなく，たくさんの局所最適解がまばらに存在するものであると仮定できる [8]．すると，打ち切り型の探索を行う際にも，探索の初期解を解空間上疎らに生成することができれば，同一の入力からでも，様々な出力が得られると考えられる．

しかし，乱数的に初期解を生成しては，最適でないような局所解に探索が落ち着いてしまうという危険を排除できない．そこで本研究では，既存の楽曲から，旋律の推移の特徴を音高の移動として形式的に記述した，ヒューリスティック・ルールベースを構築し，探索の始点の決定に利用するモデルを考えた．これらのルールには，和声学等の音楽学知識は直接実装されておらず，音高を上げる・下げる等といった機械的な操作しか含まれていない．しかしその中でも，既存の楽曲の表層的特徴を応用したこのようなルールベースを，確率モデルを用いて入力の楽譜に適用して生成された初期解を起点として探索を行うことにより，ランダムな始点から探索を開始した場合に比べて，より良い出力結果が得られると期待した．

ACC のシステム概要を，図 3.1 に示す．

[8] 芸大和声 [16] は各和声課題に対する模範解答を提示しているが，それらはあくまでも「模範的」な解答であり，正解は他にも存在するとも説明している．実際，芸大和声内でも，一つの課題に対する模範解答を複数提示している場合もある．このことからも，作曲空間における評価関数は単一の最適解を持つのではなく，局所最適解が多数存在するのだと仮定できる．ただし，これに関する具体的な解析と厳密な検証は，今後必要であろう．

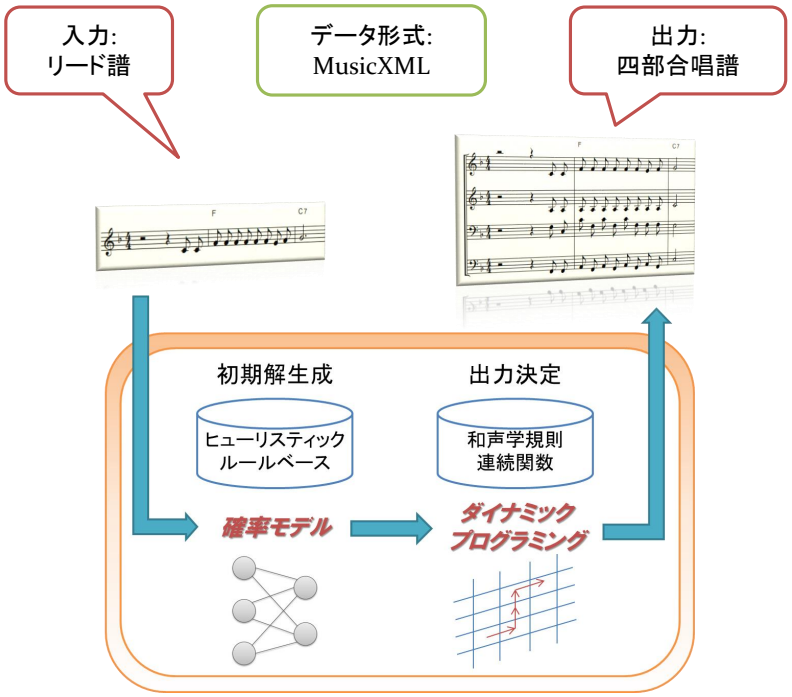

図 3.1: *Auto Chorus Creator (ACC)* システム概要図

入出力

ACC は，リード譜を入力として受け，それに基づいた 4 パートの合唱譜を出力する．リード譜は，単旋律の楽譜にコード記号を振ったものであり，多くのポピュラー音楽において利用されている楽譜様式である．作曲を始めるユーザは，単旋律の作成から始める人が多いと考えられるので，この楽譜形式を利用した．図 3.2 にリード譜の例を示す．
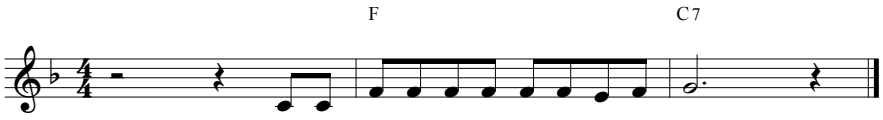
図 3.2: リード譜の例 （『If You're Happy and You Know It』の冒頭 3 小節分）

出力はデフォルト的に「合唱譜」としているが，出力の各声部の音域を変更することにより，声楽だけではなく，様々な器楽に合わせた伴奏作成も可能である．また，便宜上，入力の旋律は自動的にオクターブを変更し，出力の最高音域をもつ声部（合唱譜の場合はソプラノ声部）に割り当てている．これは，今後システムを公開に向けて拡張する際に，自由に変更可能にする予定である．また，出力の楽器数は今は 4（四声部）に限定しているが，今後その制限も解除し，任意の楽器数に合わせた伴奏作成を可能にする予定である．

入出力の楽譜は，MusicXML 形式とした．MusicXML 形式を利用した理由については，

先に紹介したシステム *CMY* で採用した理由と同じであるので，説明は第 2.2.3 節に代えさせていただく．あくまで，MusicXML 形式が市販の記譜ソフトで高い汎用性を持ち，他のソフトウェアを用いた入力の作成や出力の編集が容易であるから利用しているのであり，実際にユーザが生の MusicXML 形式のデータに触る必要はないという点は強調しておきたい．

システムは Java 言語で構築し，凡そ 5,500 行のコーディングを行った．その内，3,800 行（凡そ 70%）がシステムとアルゴリズムを含めたエンジン部分に，900 行（凡そ 15%）が MusicXML の出力生成部分に，そして 800 行（凡そ 15%）がダイナミック・プログラミングに利用される評価関数（和声学規則等）の記述に充てられている．本研究の趣旨は作曲機構そのものの提案と実装にあるため，現段階でシステムの実行は，全てコーディング内から行うようになっている．インタフェースの構築等については，今後公開に向けて GUI を設計する中で検討する予定である．

3.2 設計詳細

前節で説明したように，*ACC* は 2 つのモジュールから成り立っている（図 3.1 参照）．1 つ目のモジュールには，旋律の音高推移に関するヒューリスティックな規則を集めたデータベースと，それを用いた確率的モデルを実装し，次のモジュールの探索の始点となる楽譜を生成する．このモジュールの出力楽譜を「初期解」と呼ぶ．2 つ目のモジュールは，この初期解を入力として受け，そこからダイナミック・プログラミングを通じて最終的な出力を決定する．ダイナミック・プログラミングに用いる評価関数には，連続関数として再定義した和声学のルールベース等も実装し，それらルールの重みを調節することにより，出力楽曲の音楽特徴等が調整できる．

以下に，この 2 つのモジュールの詳細について，具体的に説明する．なお，これ以降，それぞれのモジュールは,「モジュール 1」,「モジュール 2」と呼ぶこととする．

3.2.1 ヒューリスティック・ルールベースによる初期解の生成

ACC は入力されたリード譜に対して，確率モデルによって「初期解」を生成する．初期解は，ヒューリスティック的に集められたルールベースを，設定されたパラメータに基づいて，乱数的に適用することによって作られる．このモジュールでは探索や解の評価等が一切行われないため，数百小節の楽曲についてもほぼ瞬時に初期解を生成することが可能である．

ヒューリスティック・ルールベースは，既存の楽曲に見受けられる旋律の音高の推移を元に，著者が独自に構築したものを利用した．入力楽譜の音符から，和声学等の知識は導入せず，いわば「機械的」にリズムや音高に単純な変更を施しながら，他のパートの音符を定める．その際，各変更規則 n は随所随所で確率 $p_n(0 \leq p_n \leq 1)$ で適用されることとした．この確率 p_n にはデフォルトとして初期値を与えているが，今後，ユーザが任意に変更可能にするといった実装も考えられる．これら p_n を,「編曲パラメータ」と呼ぶ．

以下に，伴奏譜 2 小節分を作曲する様子を，いくつかのルールを例に挙げて説明する．

1. 全ての小節について，それぞれのパートの音域の中から，その小節における入力楽

譜のコードの構成音を 1 つランダムで選び，全音符[9] として配置する．

2. 連続する 2 小節の音程を比較し，同じ音程が連続する場合は，後続する音程を，コードを構成する他の音程に変更する（図 3.3 参照）．

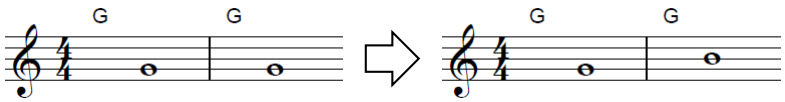

図 3.3: ヒューリスティック・ルールの例 1

3. 停止条件が満たされるまで，ヒューリスティック・ルールを適用し，各小節内の音数を変更し続ける．ここでは,「連続する 2 音が 3 度離れていて，後続音の音程が先行音より高い場合，先行音を長さの等しい 2 音に分離し，その 2 音目の音程を 2 度上げる」を適用したとする（図 3.4 参照）．

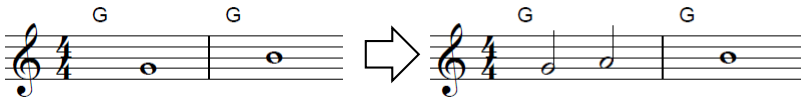

図 3.4: ヒューリスティック・ルールの例 2

このように，ヒューリスティック・ルールベースは音符を直感的に移動・変更することにより，伴奏の各パートを作成していく．そのルールには,「コード構成音」と「長さを等分する」以外に，和声やリズム等に関する明確な音楽学知識は含まれていない．モジュール 1 では,「伴奏の各パートの音数が，全ての小節において，入力の音数と同じになる」を停止条件とし，それが満たされるまでまで編集を続けるようにした．入力楽譜で 1 音も存在しない全休符の小節は，伴奏の各パートも全休符とし，入力音が存在する小節については，入力と同じだけの音数ができるまでその小節における音の追加や音高の変化を行っていく．

以下に，実装したヒューリスティック・ルール 9 個を列挙する．このルールベースは完全なものであるとは考えていない．むしろ，本研究で提案する作曲機構（2 つのアルゴリズムの併用等）が自動作曲において好ましい成果を上げるかどうか調べるために，最低限必要だと考えた数だけのルールを実装した．9 というルール数にはそれ以上の根拠はなく，今後システムを拡張するにあたり，より多様な初期解を生成できるためには，更に多くのルールを追加すべきであると考えている．

1. 同じ音程の音が連続する場合，先行音を長さの等しい 2 音に分離し，その 2 音目をコードの次の構成音に上げる．

2. 同じ音程の音が連続する場合，先行音を長さの等しい 2 音に分離し，その 2 音目をコードの次の構成音に下げる．

[9] 1 小節の時間的長さと等しい長さを持つ音符

3. 同じ音程の音が連続する場合，先行音を長さの等しい 2 音に分離し，その 2 音目をさらに 2 分する．そのうえで，2 音目は音程を 3 度，3 音目は音程を 2 度上げる.（結果として，4 音の音程は低・高・中・低と遷移する.）

4. 連続する 2 音が 3 度離れていて，後続音の音程が先行音より高い場合，先行音を長さの等しい 2 音に分離し，その 2 音目の音程を 2 度上げる.（図 3.4 参照）

5. 連続する 2 音が 3 度離れていて，後続音の音程が先行音より低い場合，先行音を長さの等しい 2 音に分離し，その 2 音目の音程を 2 度下げる.

6. 連続する 2 音が 4 度離れていて，後続音の音程が先行音より高い場合，先行音を長さの等しい 2 音に分離し，その 2 音目をさらに 2 分する．そのうえで，2 音目は音程を 2 度，3 音目は音程を 3 度上げる.（結果として，4 音の音程は低・中・高・最高と階段状に遷移する.）

7. 連続する 2 音が 4 度離れていて，後続音の音程が先行音より低い場合，先行音を長さの等しい 2 音に分離し，その 2 音目をさらに 2 分する．そのうえで，2 音目は音程を 2 度，3 音目は音程を 3 度下げる.（結果として，4 音の音程は高・中・低・最低と階段状に遷移する.）

8. 連続する 2 音が 5 度離れていて，後続音の音程が先行音より高い場合，先行音を長さの等しい 4 音に分離する．そのうえで，2 音目は音程を 2 度，3 音目は音程を 3 度，4 音目は音程を 4 度上げる.（結果として，4 音の音程は最低・低・中・高・最高と階段状に遷移する.）

9. 連続する 2 音が 5 度離れていて，後続音の音程が先行音より低い場合，先行音を長さの等しい 4 音に分離する．そのうえで，2 音目は音程を 2 度，3 音目は音程を 3 度，4 音目は音程を 4 度下げる.（結果として，4 音の音程は最高・高・中・低・最低と階段状に遷移する.）

なお，これ以外にも,「初期のコード構成音選択時に，同音程の音を連続して配置したならば，後続音を，コードを構成する別の音程に変更する.（図 3.3 参照）」というルールも実装している．このルールは，同じ音高が曲の中で過度に連続し，単調な曲が生成されないために導入した.

3.2.2 解空間における探索と評価

ダイナミック・プログラミングと探索時間

ACC はモジュール 2 において，モジュール 1 が作成した初期解を受け取り，それを始点として探索を始める．初期解は既に 4 部の合唱譜（等，ユーザが設定した編成の楽譜）として完成しているが，その音の決定において，1 つ目のモジュールにあるヒューリスティックな規則しか適応されていない．和声学やその他の音楽学に基づいた曲の構築や評価は行われていない．このモジュールでは，ダイナミック・プログラミング（以下 DP）の手法を用いて，最終的に出力する解を決定する.

DP は，取り扱う問題が複数の小さな問題に分割され，それぞれの問題の最適解の組み合わせが，当初の大きな問題の最適解を生じるような場合に有用な計算アルゴリズムである [44]．本研究では，楽曲全体の最適解が楽曲の部分部分における最適解の組み合わせから成り立つと仮定して，このアルゴリズムを実装した．第 2.1.1 節でも紹介したように，DP は他の自動作曲研究でも利用され，ある一定の精度を出していることから，本研究においても有用であると考えた．特に，DP は計算時間が膨大になりうる問題において，比較的短時間で解を求められるという特徴があるため，計算時間の理論値が指数オーダーになる本システムに実装した．

本システムにおいて，DP は，各時点における全てのパートの音からなる「和音」を評価する．伴奏パート（4 部合唱譜ではソプラノ以外の 3 パート）の音程をそれぞれ上下に一定距離移動させたうえで，最も評価値の高い和音を探索していく．DP は和声学に記載されている規則に基づき，和音の形（各パート間にどれだけの隔たりがあるか）や連続する和音の連結関係を評価していく．探索は曲の第 1 和音から順番に行われ，一度決まった和音を再度変更するようなバックトラックは実行されない．

ACC が楽曲を 1 つ作曲するのにかける時間は，このモジュールの DP 実行時間に大きく起因する．一般に作曲の解空間は指数的に広がる（第 2.1.2 節参照）ものの、DP を利用することにより、探索時間は以下のように多項式時間内に留めることが可能である。

探索の中で，和音の構成音の音程を上下にそれぞれ d 音動かすとすると，1 時刻 1 パートにつき，

$$2d + 1 \tag{3.1}$$

音を探索することになる．また，最終的に出力する楽譜のパート数を a とすると，ある時刻における探索和音数は

$$(a - 1)(2d + 1) \tag{3.2}$$

と表わせる．先の式に $a - 1$ を掛けているのは，入力の 1 パートは変更しないという仮定をしているからである．この探索を，楽曲の長さ t において実行するが，その際，1 時刻前の探索は次の探索に影響しないので，総探索和音数は

$$t(a - 1)(2d + 1) \tag{3.3}$$

と積の形で表せる．ここで，楽曲の長さ t は，全てのパートがそのすべての音を同時に演奏する場合には，入力の音数 n と等しい値になる．しかし一般に，楽曲は旋律と伴奏のリズムが異なっているため，各パートの音が重なる最短時間 t_{min} を時刻 t の単位時間としなければいけない．この最短時間 t_{min} は，楽曲に登場する最短演奏時間を持つ音符の長さである場合が多い．DP の実行時間よりも理論的に短い第 1 モジュールや出力生成等に係る時間を加味するためにオーダーの記号 $O(x)$ を使うと，結果として，システムの実行時間は以下の式で与えられる．

$$O\{\frac{n}{t_{min}}(a - 1)(2d + 1)\} \tag{3.4}$$

この中で，パート数 a と音数 n はユーザによって与えられる入力で決まるため，システム側では調節できない．よって，最短の音符の長さ t_{min} と探索の際に各音を移動させる距離 d とをシステム側で調節し，探索時間を変更させる．*ACC* では，最短の音符を 32 分

音符に設定することにより，t_{min} を定数値としている．しかし，テンポの速い楽曲ではこれをより短くする必要が生じる可能性があり，逆に遅い楽曲ではそれをそこまで細かく設定しなくてもよい可能性がある．また，ある 1 曲の中でも，旋律に合わせて t_{min} を変動させることにより，無駄なく計算時間を抑えられそうである．今後の実装として，入力の楽譜に基づいて t_{min} を可変的に決定する機構が考えられる．

結論として，DP を利用することにより，ACC の実行時間は，探索の距離 d によってのみ決まること，またそれが多項式時間内に終了することが分かる．

探索における評価関数

上述したように，ACC に実装した DP の評価関数には和声学の諸規則が実装されている．しかし，和声学の規則は「これこれしてはいけない」等と，絶対的な記述がされていて，そのまま計算機に実装すると，$true$ か $false$ の 2 値を返す論理関数の形になる．論理関数は計算しやすく，和声学に反するか反しないか等と楽曲を判定する作業においては有用である．しかし，同等に優れている（劣っている）楽曲を仕分ける操作には不向きである．そのため，多くの先行研究において，禁則を実装した評価関数には実験者の主観による重みが割り当てられ，各規則を破った回数の比較によって解の評価を行っている [7]．三浦らは，各禁則の重み（加減点の比率）を他の作曲課題の評価から算出する評価システムを構築している [25] が，規則違反に対する 2 値処理を行っていることには変わりない．

規則違反を 2 値的に観察すると，ある一つの規則に異なった方法で違反している 2 つの楽曲の評価に差を設けることができない．異なっている楽曲にも同じ評価がついてしまうと，探索における優劣を別な要素から定めなければいけなくなり，評価機構が助長的，あるいは主観的になり兼ねない．そこで本研究では，各規則に違反の度合いを考慮した連続な評価関数を定めることにより，実質的に異なった楽曲にはなるべく異なった評価値がつくようにシステムを設計した．和声学の諸規則に連続な評価関数 [10] を定めた先行研究は存在しなかったため，ACC には著者が独自で再定義した評価関数を割り当てた．以下に，芸大和声 [16] から用いた各規則の原文と，その実装をするにあたって設定した評価関数を記述する．各評価関数の番号は，芸大和声の規則番号に対応する．また，評価関数内の変数については，表 3.1 に説明を記した．なお，規則分の中に和声学の専門用語が頻出するが，これらについては直接 [16] を参照されたい．

なお，探索は最小化問題として実装している．関数はどれも 0 から 1 の間の値をとり，良い場合には 0 を，悪い場合には 1 が算出される．ただし実際には，いくら悪い楽曲を書いても，全ての関数を同時に 1 にすることは不可能であり（証明は省く），探索における評価値も，1 をとることはない．

[10] 厳密には，音符（音高）は「ド・ド♯・レ・レ♯...」等と離散的な値をとるため，本稿で設定した関数もすべて離散値をとる．ただし，分かり易さのために，連続関数の形で各評価関数を記述することとした．

A 1 すべての和音を通じて，省略しうる構成音は第 5 音だけである．ただし，V_7 の和音・V_9 の和音にかぎり，根音を省略することがある

$$f_{A1} = \begin{cases} \dfrac{n_{abbr}}{n_{chord} - 2} & (chord = V_7\text{の和音，または } V_9\text{の和音の場合}) \\ \dfrac{n_{abbr}}{n_{chord} - 1} & (\text{それ以外の場合}) \end{cases} \tag{3.5}$$

A 2 限定進行音を重複してはならない．

$$f_{A2} = \frac{n_{leading} + n_{7th}}{a - 1} \tag{3.6}$$

A 3 V_9 の和音の第 9 音は，根音より 9 度以上上方に置かなければならない．

$$f_{A3} = \frac{|min(p_{9th} - p_{tonic} - 14,\ 0)|}{(p^{max}_{a_{tonic}} + 14) - p^{min}_{a_{9th}}} \tag{3.7}$$

A 4 長調の V_9 の和音の第 9 音は，第 3 音より 7 度以上上方に置かなければならない．

$$f_{A4} = \begin{cases} \dfrac{|min(p_{9th} - p_{3rd} - 10,\ 0)|}{(p^{max}_{a_{3rd}} + 10) - p^{min}_{a_{9th}}} & (\text{長調の場合}) \\ 0\ (\text{それ以外の場合}) \end{cases} \tag{3.8}$$

A 5 変位構成音と他の構成音とが減 3 度（単音程）をなしてはならない．

$$f_{A5} = \frac{s_{p_{accidental}, p_{other}}}{n_{p_{accidental}} \times (a - 1)} \tag{3.9}$$

B 1 進行において制限される音程はつぎのとおりである．
長・短 7 度
増 1 度を除く増音程
複音程

$$f_{B1} = \frac{\sum_a g_{restricted}}{a} \tag{3.10}$$

$$g_{restricted} = \begin{cases} 1\ (\text{制限される進行を形成している場合}) \\ 0\ (\text{その他の場合}) \end{cases} \tag{3.11}$$

B 2 限定進行音は，所定の限定進行（2 度上行・下行）か，または保留，増 1 度進行のいずれかをなさなければならない．

$$f_{B2} = \frac{\sum_a g_{tritone}}{a} \tag{3.12}$$

$$g_{tritone} = \begin{cases} \dfrac{|p_{a_{tritone}}(t) - p_{a_{tritone}}(t+1) - 1|}{p^{max}_{a_{tritone}} - p^{min}_{a_{tritone}}} \\ (\text{所定の進行を行っていない場合}) \\ 0\ (\text{その他の場合}) \end{cases} \tag{3.13}$$

B 3 連結すべき 2 個の和音の構成音のうちに半音階的関係をなす 2 音が含まれる場合には，それらの 2 音は，ある 1 個の声部で増 1 度関係に連結しなければならない．

$$f_{B3} = \frac{\sum_{a \cdot a} g_{semitone}}{a^2} \tag{3.14}$$

$$g_{semitone} = \begin{cases} \dfrac{|p_{a_{semitone}}(t) - p_{a_{semitone}}(t+1) - 1|}{p^{max}_{a_{semitone}} - p^{min}_{a_{semitone}}} \\ （同一声部内に半音階的関係をなす 2 音がある場合） \\ \dfrac{|p_{a_{semitone},a_1}(t) - p_{a_{semitone},a_2}(t+1) - 1|}{max(p^{max}_{a_{semitone},a_1},\ p^{max}_{a_{semitone},a_2}) - min(p^{min}_{a_{semitone},a_1},\ p^{min}_{a_{semitone},a_2})} \\ （異なる声部内に半音階的関係をなす 2 音がある場合） \\ 0（それ以外の場合） \end{cases} \tag{3.15}$$

C 1 連続 8 度・連続 1 度は禁ぜられる．

$$f_{C1} = \frac{\sum_{a \times a} h_{restricted_{C1}}}{{}_aC_2} \tag{3.16}$$

$$h_{restricted_{C1}} = \begin{cases} 1（規則\ C1\ に反する場合） \\ 0（その他の場合） \end{cases} \tag{3.17}$$

C 2 連続 5 度は，後続音程が完全 5 度の場合に禁ぜられる．

$$f_{C2} = \frac{\sum_{a \times a} h_{restricted_{C2}}}{{}_aC_2} \tag{3.18}$$

$$h_{restricted_{C2}} = \begin{cases} 1（規則\ C2\ に反する場合） \\ 0（その他の場合） \end{cases} \tag{3.19}$$

C 3 並達 8 度・並達 5 度は，外声間に生じ，かつ，ソプラノの跳躍進行による場合にかぎり禁ぜられる．

$$f_{C3} = \frac{\sum_{a \times a} h_{restricted_{C3}}}{{}_aC_2} \tag{3.20}$$

$$h_{restricted_{C3}} = \begin{cases} 1（規則\ C3\ に反する場合） \\ 0（その他の場合） \end{cases} \tag{3.21}$$

C 4 並達 1 度は禁ぜられる．

$$f_{C4} = \frac{\sum_{a \times a} h_{restricted_{C4}}}{{}_aC_2} \tag{3.22}$$

$$h_{restricted_{C4}} = \begin{cases} 1（規則\ C4\ に反する場合） \\ 0（その他の場合） \end{cases} \tag{3.23}$$

C 5 第 7 音の 2 度下行と根音の 3 度下行とによる並達 8 度はいずれの声部間においても禁ぜられる．

$$f_{C5} = \frac{\sum_{a \times a} h_{restricted_{C5}}}{{}_aC_2} \tag{3.24}$$

$$h_{restricted_{C5}} = \begin{cases} 1\text{（規則 } C5 \text{ に反する場合）} \\ 0\text{（その他の場合）} \end{cases} \tag{3.25}$$

探索における評価関数として，他に和声学を利用しない関数も導入した．探索の各評価関数は重み付けされ，その重みをユーザによって変更可能ともできる．しかし，ACC がターゲットとしている「作曲に関する専門知識を有しないユーザ」には，これら評価関数の理解やパラメータの細かい変更が難しいと考えた．そこで，専門知識を有しないユーザにも理解できる，より一般的な特徴量「難易度」を導入し，それも探索において評価することとした．

一般に，ある楽曲の難易度を評価する統一的な手法はない．そもそも，楽曲の難易度をどう定義するかというのも，音楽家によってまちまちである．本研究ではとりあえず，「パート内での音高の遷移量」と「パート内の非和声音の数」という 2 つの楽曲特徴を定義し，それぞれが多いと楽曲の難易度が高くなるとした．これらの具体的な評価関数を以下に提示する．なお，Z1 は全パートについて，Z2 は全和音について更に総和が取られ，規格化されて，探索の評価に用いられる．評価関数内の各変数については，表 3.1 に併せて説明を記した．

Z 1 音高遷移量 ある声部内での合計音高遷移量

$$f_{Z1} = \frac{\sum_{a_x} \frac{||p_{a_x}(t) - p_{a_x}(t+1)| - 2|}{p_{a_x}^{max} - p_{a_x}^{min} + 1}}{N_{a_x}} \tag{3.26}$$

Z 2 非和声音数 ある和音における非和声音量

$$f_{Z2} = 0.2 \times \{i(p_{soprano}) + i(p_{alto}) + i(p_{tenor})\} + 0.4 \times i(p_{bass}) \tag{3.27}$$

$$i(p_{a_x}) = \begin{cases} 1\text{（}p_{a_x}\text{が非和声音の場合）} \\ 0\text{（その他の場合）} \end{cases} \tag{3.28}$$

評価関数 f_{Z1} の分子の分子は，設計当初は $|p_{a_x}(t) - p_{a_x}(t+1)|$ とし，同じ音程が続いたときに評価値が最も小さくなるようにしていた．しかし，著者の実体験から，実際に楽器を演奏したり声楽で歌ったりするときには，長時間同じ音を出し続けることは意外と難しいと感じるのである．特に，音の細かい音程にこだわりを持つようになると，2 つの異なる音程が連続する方が，同じ音程を連続して演奏する場合よりも，その音程の誤差が聞こえにくい傾向がある．これを踏まえ，楽曲の「難易度」としてこの特徴量を導入するにあたり，旋律が少し変動するときにもっとも評価値が低くなるように，上述したように 2 を引いている．

また，評価関数 f_{Z2} についてだが，バスパートとそれ以外のパートとにおいて，重みは意図的に変更している．通常，ある楽団の最低音楽器は，楽曲のコードを演奏する風潮が

ある．クラシック音楽におけるティンパニーやチューバは特にそうであり，ジャズ音楽のベース，バンド音楽のベースギター等もそうである．そのため，バスパートが非和声音を奏でることは，総合的に楽曲を難しくするだろうと仮定し，評価関数においても，それを他のパートよりも重点的に評価している．

3.2.3　ユーザ視点でのシステム概要

以上のように，*ACC* は 2 つのモジュールからなり，内部には様々な変数が設定されている．以下に，ユーザ視点からのシステム概要をまとめ，その様子を図 3.5 に示す．

ユーザは,「リード譜」と「難易度」を入力する．リード譜は MusicXML 形式で入力されるが，これは市販の記譜ソフトでも作成可能である．難易度は 0 から 1 の間で与える．現在は作曲の複雑さを示す一つの指標として導入しているが，一般的に言われる「曲の難易度」との対応関係については，今後検証を必要とする．

ユーザは他に,「各パートの音域」や,「評価に用いる和声規則の指定 」等のパラメータも入力することができる．しかし，これらはそれぞれシステム内にデフォルト値を持っていて，システムの実行体系には大きく影響しない．

また，最終的な出力として，ユーザは 4 パート全てが並記された楽譜（スコア譜）か，それぞれのパートが個別の楽譜として記載されるパート譜の 4 枚組のいずれかを選択できる．両方の出力も可能である．

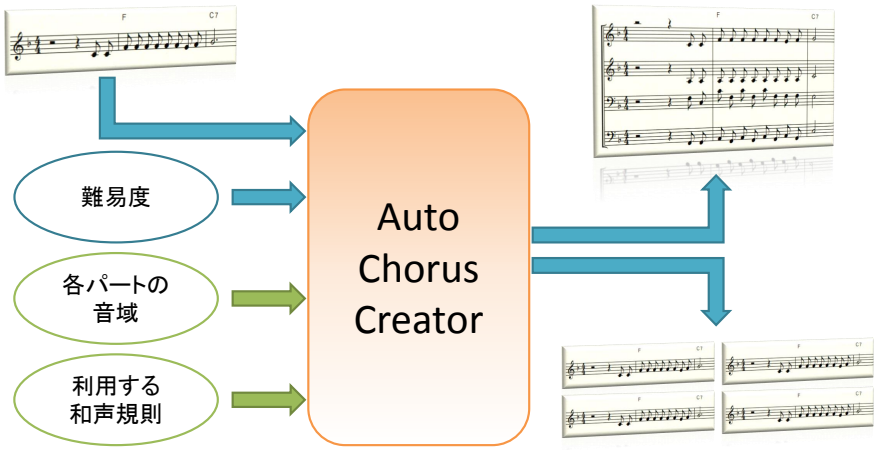

図 3.5: 自動作曲システム *ACC* のユーザ視点での概要図

表 3.1: *ACC* 第 2 モジュールの評価関数に用いた各変数の説明

a	楽曲の総パート数
$\sum_a$	全パートにおける総和
$\sum_{a \cdot a}$	重複ありの全パート総組み合わせ数
$\sum_{a \times a}$	重複なしの全パート総組み合わせ数
n_{7th}	和音内の第 7 音の数
n_{abbr}	省略してはならないもののうち，省略されている和音構成音の数
n_{chord}	和音の構成音の数
$n_{leading}$	和音内の導音の数
$n_{p_{accidental}}$	和音内の変位構成音の数
p_{3rd}	和音内の第 3 音の音高
p_{9th}	和音内の第 9 音の音高
p_{tonic}	和音内の根音の音高
$p_{a_{tritone}}(t)$	時刻 t の和音において，限定進行音を行っている声部の音高
$p_{a_{semitone}}(t)$	時刻 t の和音において，半音階的関係をなしている声部の音高
$p_{a_{semitone},a_x}(t)$	時刻 t の和音において，半音階的関係をなしている x 番目の声部の音高
$p^{max}_{a_{3rd}}$	和音内で第 3 音を持っている声部の声域内最高音の音高
$p^{min}_{a_{9th}}$	和音内で第 9 音を持っている声部の声域内最低音の音高
$p^{max}_{a_{tonic}}$	和音内で根音を持っている声部の声域内最高音の音高
$p^{max}_{a_{semitone}}$	半音階的関係をなしている声部の声域内最高音の音高
$p^{min}_{a_{semitone}}$	半音階的関係をなしている声部の声域内最低音の音高
$p^{max}_{a_{semitone,x}}$	半音階的関係をなしている x 番目の声部の声域内最高音の音高
$p^{min}_{a_{semitone,x}}$	半音階的関係をなしている x 番目の声部の声域内最低音の音高
$p^{max}_{a_{tritone}}$	限定進行を行っている声部の声域内最高音の音高
$p^{min}_{a_{tritone}}$	限定進行を行っている声部の声域内最低音の音高
$s_{p_{accidental},p_{other}}$	和音内で変位構成音と他の構成音とが減 3 度をなしている箇所の数
$\sum_{a_x}$	パート a_x における総和
N_{a_x}	パート a_x の総音数
$p_{a_x}(t)$	時刻 t の和音における，パート a_x の音高
$p^{max}_{a_x}$	パート a_x の声域内最高音の音高
$p^{min}_{a_x}$	パート a_x の声域内最低音の音高

第4章　システム評価と結果

本章では，実装したシステムとその出力に行った評価について説明する．システムは2つのモジュールから成り立っているため，まず，それらを個別に用いた場合の性能評価を説明する．その後，システム全体としての評価を，実際に出力された楽譜等を例にあげてまとめる．

4.1　モジュールごとの評価と考察

モジュール1と初期解の生成

第1モジュールは，様々な初期解を生成し，第2モジュールのDPが違う起点から探索を開始できるようにする目的で実装された．そこで，第1モジュールは，同じ入力楽譜から異なる初期解を生成されているかどうか，という視点で評価する．

図4.1に，モジュール1の出力例を3つあげる．3つの出力はどれも，最上部に記載したソプラノ音とコード譜を入力として与えた．入力は，図3.2にも示した，2007年に文化庁により『日本の歌百選』にも指定された「幸せなら手をたたこう」（英語名：”If You’re Happy and you Know It”）を用いた．紙面の関係上，ここには冒頭4小節のみを記載する．それぞれの初期解を生成するにあたり，ヒューリスティック・ルールベースや確率モデルの各パラメータ値は同じにし，システムの乱数の種だけを変更した．

ヒューリスティック・ルールベースは，各パートの各小節における音数がソプラノと同じになるまで適用され続ける．また、数式3.4における最短の音符長t_{min}は，ここでは，入力譜の中で最も短い八分音符の長さとしている．この楽曲は，1小節が八分音符8個で埋められるため，最大の音符数で埋められている2小節目及び4小節目には，各パート間でリズムの違いは生じていない．ただし，第3小節目の出力例を見て分かる通り，小節の中での音符の割り振り方に自由度があるとき，システムは異なったリズムをパート間で作成することができている．

また，出力1と出力3のバス譜（最下部のパート）の最後3音を見てみると,「ド→レ→ミ」と階段状の作曲がされている．出力2ではそれは行われていない．これはヒューリスティック・ルールの中に，階段状に音を配置するルールがあるために起きている現象であるが，この3音の中でも「レ」の音は，その時の和音（C^{7th}のコード）を構成する音に含まれない，いわゆる非和声音となっている．この非和声音は，コードの移行時等に頻繁に用いられる傾向がある．第3.2.1節で記載したヒューリスティック・ルールには，コードが変化するときにこの階段状の移行を利用せよ，等といった特別な記述はない．しかし，このヒューリスティック・ルールベースに実装した単純な構造的なルールでも，実際に用いられるような作曲スタイルを適用することができている.（他にも，出力1のテナーパート
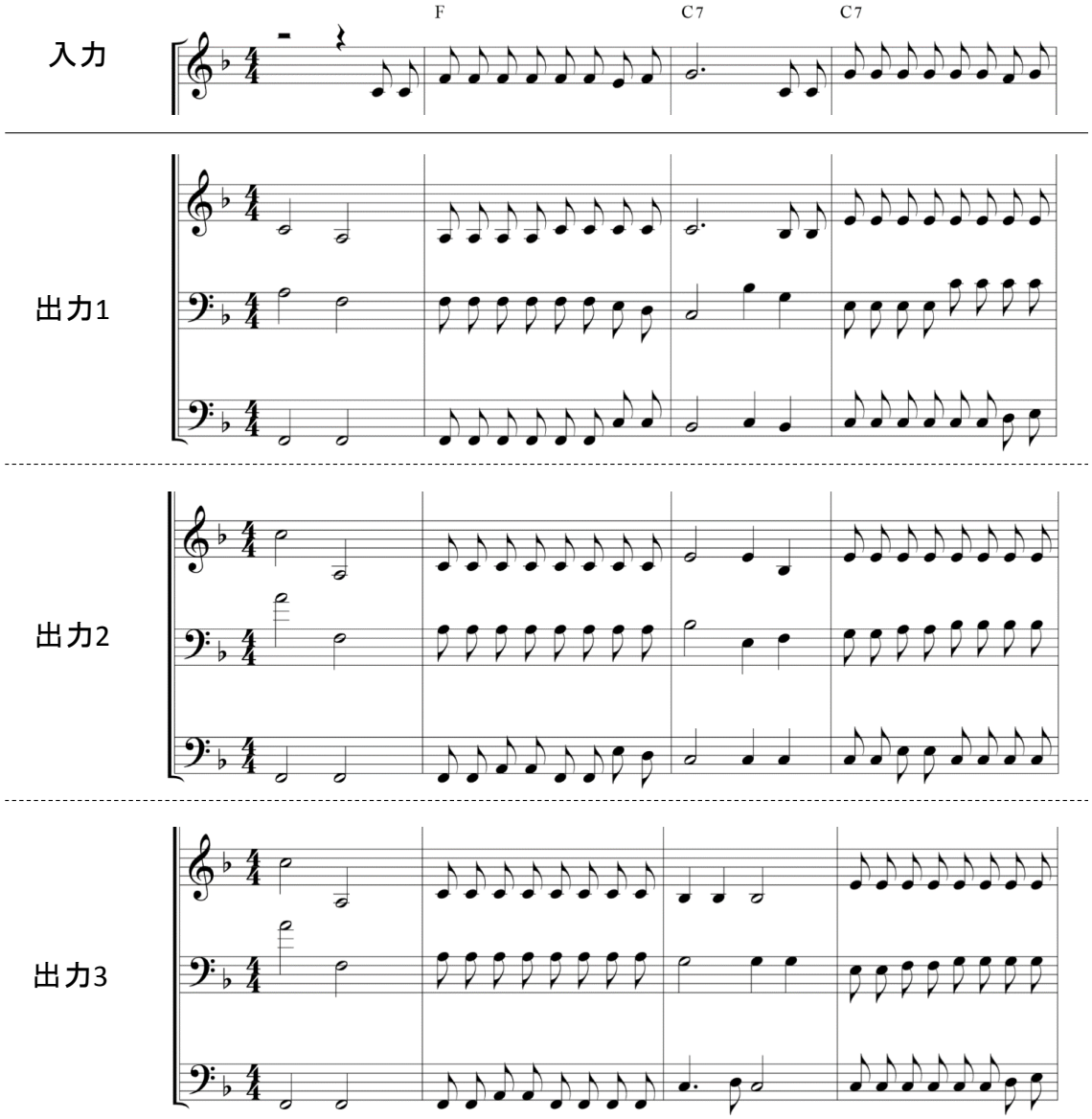
図 4.1: 自動作曲システム *ACC* のモジュール 1 のみ利用した際の出力例（それぞれ，上からアルト，テナー，バスのパートの出力例を表わす）

の 2 小節目の最後，出力 2 のバスパートの 2 小節目の最後等にも同じ現象が起きている.）

一見似ているように見えるこれら伴奏も，よくよく見るとその音高は様々な場所で異なっていることが分かる．全 21 音中，出力 1 と出力 2 とで同じ位置で同じ音高を持つ音は 1 音，出力 1 と出力 3 とでは 4 音，出力 2 と出力 3 とでは 10 音しか存在しない．ランダムに生成したこの 3 つの出力でも，半分以上同じ音が登場しているものはない．

更に，これら出力は和声学の規則を破っていて，和声的にはまだ改善の余地があることにも気が付く．特に出力 1 と出力 2 の冒頭のアルトとテナーのパートにおいて，和声学では禁じられている大きな跳躍進行が起きていることが目立つ（第 3.2.2 節の和声学規則 B1 を参照）．この和跳躍進行は，和声学的に良くないだけではなく，実装に歌う（演奏する）ことも困難な，高度な旋律の形である．また，いずれの出力において，3 小節目最後の限定進行音「シ♭」が，所定の限定進行を行っていない（第 3.2.2 節の和声学規則 B2 を参照）のも，和声学的には問題である．

このように，単純なヒューリスティック・ルールの確率的適用により，異なる初期解は確かに生成されていることを確認した．しかし，それぞれの初期解の中には，和声学的に改善の余地を残す部分や，演奏難易度が非常に高い部分も存在する．これらの部分の訂正，及び総合的に質の高い楽曲生成の保証は，次のモジュールに任せることになる．

モジュール 2 と出力の生成

モジュール 2 は，モジュール 1 で生成された初期解を起点として，各音の音高を上下に距離 d だけ動かして，評価値の高い出力を探索する．そこで，モジュール 2 は，与えられた初期解の近傍で，更に評価の高い出力を正しく探索できているか，という視点で評価する．

上に倣い，図 4.2 に第 2 モジュールの出力例を 3 つあげる．どの出力も，図 4.1 の初期解起点とした，DP の出力結果である．探索には，第 3.2.2 節で提示した規則をすべて利用し，各パラメータにはデフォルト値を，数式 3.4 における探索の距離 d には 7 を利用した．また表 4.1 には，各出力に対して第 2 モジュールを利用した際の，各評価値の値の変化を示す．

図 4.2 の出力結果から，まず非和声音がほとんどなくなっていることに気が付く．前節で言及した，ヒューリスティック・ルールによって実現された階段状の音の移行は，全ての音が和音の構成音になるように置き換えられている．ただし，出力 3 の 4 小節目の最初のアルト音のように，初期解にはなかった非和声音があえて加えられている場合もある．非和声音が増えると，旋律が和声の諸規則を破る可能性が増えるとともに，合唱等の実際の演奏の際にも，その難易度が上がると言える．このように多くの非和声音が消えているのは，評価関数に実装された和声規則の効果による．ただし，今現在の評価関数には，コード移行の際の階段進行等のように，非和声音をあえて残すためのルールや評価関数が組み込まれていない．これは今後システムを拡張した際に追加しようと考えている．

和声学では，同じパート内で同じ音高が続く現象を避ける傾向がある．これは，第 3.2.2 節の規則群 C 等で，複数パート間での同音程の連続に関する規則等を設けることによって一部表現されている．また，同じ音が続く楽曲は「面白み」に欠けるとする教科書本文の書き方などから読み取れる．しかし，これら和声の諸規則を守りながらも，単一のパート内に同じ音程を連続させることが可能な場合も多い．そのため，著者は規則 Z1 で音高を移動
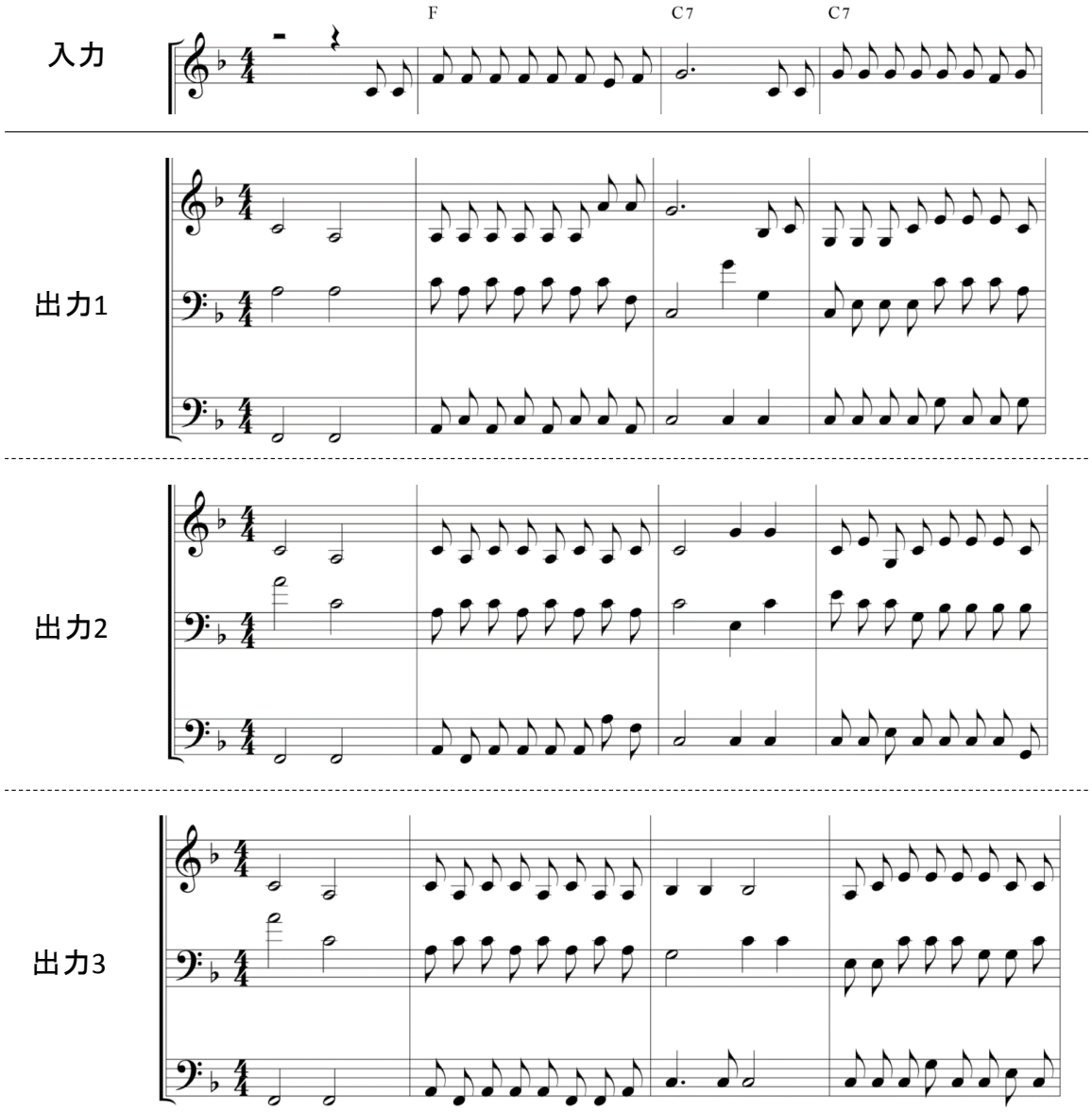
図 4.2: 自動作曲システム *ACC* のモジュール 2 を，複数の初期解に利用した際の出力例（それぞれ，上からアルト，テナー，バスのパートの出力例を表わす）

させた場合に最小値をとるような評価関数を実装している．この効果により，モジュール 2 の出力には，同じパート内で同じ音高が続く個所は極端に少なくなっている．モジュール 1 が生成する初期解には同じ音高が 1 小節も続く部分もあったが，それはこれら出力ではなくなった．

他にも，限定進行音が所定の進行を行っていなかった部分は訂正されている点（出力 1 及び出力 3 のアルトの 3 小節目最後の音）や，出力 2・出力 3 の冒頭におけるテナーの極端な跳躍進行が訂正されている点等から，DP を用いたモジュール 2 の出力は，全体的に，和声学的により整った出力になっていることが言える．ただし，その出力は，和声学的に完全であるというわけではない．出力 1 のテナーの 3 小節目には，初期解にはなかった極端な跳躍進行が出現している．DP は総合的な評価値を考慮して，一部の規則に従わない箇所を残す選択も行っている．

モジュール 2 によって出力された最終的な楽譜が，モジュール 1 によって生成された初期解の楽譜よりも，評価関数の各規則に従っているという事実は，図 4.1 の結果からも読み取れる．どの出力をどの規則群から見ても，DP 実施後の評価値は良くなっていて，それが初期解の半分以下の値になっている項目も少なくない．

表 4.1: モジュール 1 に生成された初期解（図 4.1），及びモジュール 2 に生成された出力（図 4.2）に対する，各項目の評価値の比較

	出力 1		出力 2		出力 3	
	初期解	出力	初期解	出力	初期解	出力
規則群 A	0.0806	0.0398	0.0667	0.0398	0.0759	0.0444
規則群 B	0.0356	0.0160	0.0319	0.0172	0.0224	0.0097
規則群 C	0.0469	0.0103	0.0225	0.0056	0.0254	0.0084
規則群 Z	0.0727	0.0373	0.0838	0.0303	0.0666	0.0305
合計	0.2358	0.1034	0.2049	0.0929	0.1903	0.0930

各楽曲について，各規則群の評価値を単純に足し合わせた合計値も記載したが，本稿で紹介した出力探索の際にシステムが用いた評価値は，この値である．各規則群を単体で見た際の評価値は，探索そのものには利用していない．評価関数に，これら規則群の重みづけ和を利用すると，特徴の異なった出力を生成できることも調べているが，その詳細については [48] に代えさせていただく．

出力結果の統計的解析

従来の自動作曲研究では，以上で述べたように，実験者が選んだ限られた楽曲について，少数の事例を挙げながら，その評価値の変化等を参考に，システムの評価と考察が行われる．聴衆は，提示された楽曲を頭の中で再現したり，または実際に楽器を演奏したりして楽曲を評価し，そこから間接的にシステムや作曲機構を評価する．実験者の考察と聴衆の考察が合致するときに，その実験者の発表は説得力を増すが，これら評価方法が非常に主観的なものになっている故に，実験者と聴衆とが結果を共感できない場合も少なくない．そこで本研究では，事例の比較にのみ頼らない，より客観的なシステム評価方法を検

討した．

まず，以上で用いたものと同じ楽曲（図 3.2）について，初期解の生成にヒューリスティック・ルールベースを用いた場合と用いなかった場合（モジュール 1 を用いた場合と用いなかった場合）とでの出力の評価値の変化を調べた．乱数の種を変えた際の，同じ入力に対する 100 回の出力の評価の平均値を図 4.3 に示す．青いグラフは，初期解の生成にヒューリスティック・ルールベースを利用した場合，赤いグラフは各パートの声域内からランダムに音高を選んで配置した場合を表わす．また，参考までに，各パートの音域の平均音高を連続して並べた楽曲を初期解としたときの実行結果を緑のグラフで示した．なお，この緑のグラフについては，各実行において全く同じ結果が出力されるので，一回の結果のみを記載している．

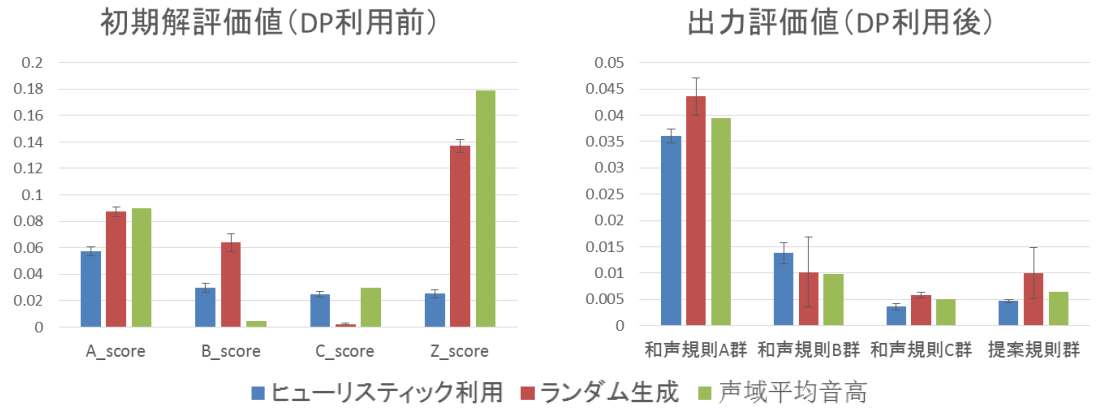

図 4.3: モジュール 1（ヒューリスティック・ルールベース）の利用の有無による初期解評価値，及び出力評価値の比較（同じ入力に対する 100 回のシステム実行の平均．エラーバーは標準偏差を示す．）

まず，図 4.3 のグラフにおいて，分散分析を行ったところ，8 項目全てにおいて，ヒューリスティック利用とランダム生成との間に強い有意差が生じた ($p \ll 0.01$)．これは DP の起点（初期解）生成に，ヒューリスティックの利用が有用であることを示している．ヒューリスティック・ルールには，和声学に関する知識は直接実装していなかった．しかし，旋律の表面的な構造のみを実装したこれらルールでも，最終的には，ランダムな起点よりも良い結果を生じている．理論的なシステム計算時間は増えていないが，このように 2 つのアルゴリズムを加えることにより，探索アルゴリズムのみを用いた場合よりも良い結果が得られたことが，客観的に示された．

和声学を直接実装していないルールを用いて生成した初期解が，ランダムに生成したものよりも，和声学的に良い結果を最終的に生んだ理由として，以下のことが考えられる．作曲を行うとき，作曲家は旋律やリズムや和音進行など，楽曲の様々な要素を順に決定・変更していくものの，その一つ一つのプロセスにおいて扱っている要素は 1 つではない．特に実際の音を楽譜に書き落として瞬間、その音は音高とリズムの 2 つの要素を同時に与えられるのであり，作曲者も総合的にそれら要素の組み合わせを考慮して評価している．あ

るフレーズ[1]を決定する際，作曲者は音高や和音やリズムのみを孤立的に扱うのではなく，前後の音楽的文脈や楽曲全体におけるポジションからそのフレーズのあるべき姿を総合的に判断し，決定する．それは旋律やリズム等，楽曲の要素を単独で扱っただけではしっかり評価できない．

つまり，音高が階段状に遷移することが多いといった旋律の特徴は，現在体系的にまとめられている和声学の概念が直接導入されてはいないものの，作曲者が総合的に楽曲を評価し，作成してきた中で培われた音楽学知識であり，その背後に和声学の影響が全くなかったとは言えないのである．音高の階段状の遷移等は，著者が既存の楽曲からヒューリスティック的に選択した楽曲の特徴であり，そのルールの実装は，数多くの実例を学習した末に決定された．和声学も，それと同じように，多くの既存の楽曲を音楽の専門家たちが解析し，似たような傾向を抽出して規則として定義されたものである．究極的には，ちょあ者のヒューリスティック・ルールベースも，体系化された和声学も，どちらも事例の学習に基づく，現象を説明する知識の集まりに過ぎず，似たような事例に基づいて形成されている以上，その生じる楽曲がある程度共通の特徴を持ったものになること（この場合，二つの知識ベースを組み合わせることにより，単独で用いるよりも良い解が得られること）は驚きではないのかもしれない．

ただし，ヒューリスティック・ルールベースの利用があらゆる面において，ランダムに生成された初期解よりも良い成果を出しているのではない．図 4.3 から分かる通り，初期解の評価値を見たときに，規則群 C の評価値に限ってはランダム生成がより良い結果を出している．規則群 C は，和音の連続に関する公理を述べた規則群になっているため，連続する和音間で，それぞれ 2 つの音を同時に評価しているものとなっている．本研究で実装したヒューリスティック・ルールベースには，ある一つの声部内で音高を調節するルールは備わっているものの，複数声部をまたいで和音を考慮するルールは一つも入れていない．そのため，規則群 C については，ヒューリスティックを利用した場合に比較的高い値が出ている[2]．しかし，ダイナミック・プログラミングによる探索後には，連続する和音に関するルールの記述がなかったヒューリスティックの方が良い結果を出しているのは興味深い．

また，探索後には，ランダムに生成した初期解を用いた方がヒューリスティック・ルールを利用した場合に比べて，規則群 B について良い結果を出している．しかし，探索前の初期解生成時の評価を見ると，その時点では，ヒューリスティック・ルールを利用した場合の方が有意に良い解を生成していた．そのため，最終的な出力時におけるこの評価値の差は，初期解の生成に直接起因しているものではないと分かる．おそらく，探索の段階における規則群の競合によって，全体の評価値に占めるそれぞれの評価関数の配分が変わり，伴奏生成に起因する各規則の影響力が変動するのだろうと考えられる．現段階でそれぞれの連続関数は，一次関数の形で書かれている．しかし厳密には，反比例やあるいは対数関数のように，グラフの変化量が変動するような関数で実装する方が，各規則に対する実際の作曲者の意識に類似しているのかもしれない．各連続関数の定義や，評価段階における諸規則の競合などについては，今後更なる検証が必要である．

[1]複数の音が連なって生成する，楽曲における「文」のようなもの．通常，楽曲は複数のフレーズの組み合わせから成り立っている．

[2]第 3.2.2 節にある通り，本研究における評価関数は最小化関数として定義されている．そのため，より小さな値を返す楽曲の方が「良い」と評価されている．

探索の評価関数はこのようにさまざまな規則によって行われるが，実際の解決定はそれらの和で決められる．図 4.4 に，複数の楽曲について，ヒューリスティックルールを利用した場合と，ランダムに初期解を生成した場合とでの，最終的な出力の評価値の違いを示す．分散分析を行った結果，これらの場合においても，ヒューリスティック・ルールベースの利用がランダムな初期解再生よりも有意に良い結果を導いていることが分かった ($p \ll 0.01$)．

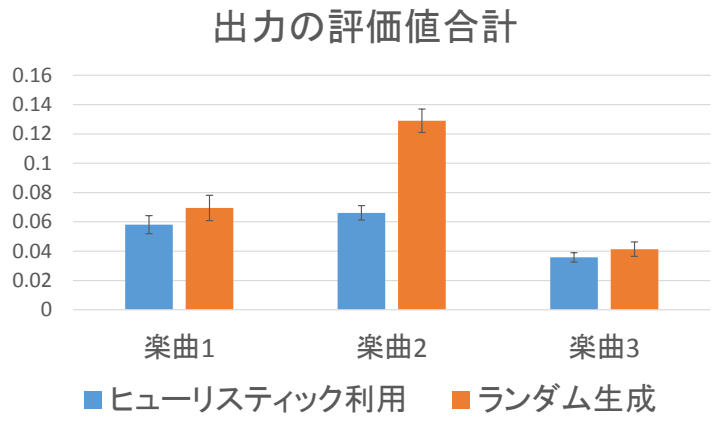

図 4.4: ヒューリスティック・ルールベースの利用の有無による，複数楽曲に対する最終的な出力の解評価値の比較（同じ入力に対する 100 回のシステム実行の平均．エラーバーは標準偏差を示す．）

4.2　システムに対する総合的な考察

第 2.4 節に，実用的な自動作曲システムに求められる特徴を 3 点述べた．そのそれぞれの点について，本研究の成果を考察する．

一点目は，音楽学の知識が，システムに内蔵されない，容易に拡張可能な形で実装されることだった．実装形態そのものから，*ACC* がその特徴を満たしていることが分かる．ヒューリスティック・ルールベースに実装されたルールには,「和声音」や「音階」という和声学上の概念は含まれていたが，その細かい知識（具体的な和音構造に関するもの等）は実装されていなかった．また探索アルゴリズムそのものにも音楽学の知識は含まれず（探索領域は和声音・非和声音の区別なく行われた），それは全てアルゴリズムから切り離された評価関数の中に実装されていた．これにより，*ACC* は今後も拡張が容易であり，和声学の他の規則，非和声音の保持に関する規則，リズムに関する規則等，他の諸規則を導入しても，アルゴリズムそのものを変更する必要がないという特徴を持っている．これは，この作曲機構を他の音楽ジャンル等に適応する際にも効果的に働くだろう．

二点目として，膨大な解空間を適切に探索し，実時間内に出力を生成可能であることをあげた．*ACC* の探索は曲の第 1 和音から順番に行われ，バックトラッキングは行われない．つまり，一度決まった和音は次の和音の評価時に参照されることはあっても，それ以上変更されることはない．また，探索の後に楽曲に何かしらの処理を施すことも行っていないため，DP の探索結果がそのまま最終的なシステムの出力となる．そのため，曲の冒

頭から，探索が終了した時点で順次その和音を出力することができる．出力された音は，再生・演奏されるのにある有限な時間を必要とする．そこで，次の和音の探索と出力にかかる時間をその有限時間以内に収めることができれば，実質的に実時間実行の自動作曲システムを構築できたことになる．特に，ACC が出力する楽曲内で最短の音の長さ（数式 3.4 の t_{min}）は既知の情報としてシステム内に保持されているので，その時間内に次の出力が生成できれば良い．旋律に合わせて t_{min} を動的に変更した場合も，それに応じて探索の距離 d も動的に変更させれば良いので，計算にかかる実時間は容易に調整可能である．このように，決定された音を随時出力する場合，MusicXML のデータ形式は最適ではないだろうが，システムの入出力形式を変更さえすれば，原理的にはアルゴリズムを変更せずに，そのまま実時間実行可能である．

三点目として，同一のシステムから，特徴が有意に異なる様々な楽曲が生成可能であることをあげた．従来の自動作曲システムでは，作曲システムのパラメータや，システムが作曲前に行う学習のデータ等を調節することにより，出力の結果を変更しているものが大半であった．本稿のシステムでも，同じようにパラメータの変更によって出力の調整を行うことは可能である．しかそれとは別に，パラメータを全く調節せずに，同等に優れているような，特徴の異なった出力を生成することも可能である．ただし，初期解の生成が乱数に基づいていることにより，システムの利用におけるユーザの自由度が制限されてしまうという実態は，このシステムの一つの問題である．今後，ユーザにある程度の自由は確保しつつも，同等に優れている複数の解の生成がユーザの細かいパラメータ調整に頼らない，バランスの取れたシステムの設計について更なる検討が必要だ．

なお，別に考慮すべき点として，システムやその生成結果が客観的に評価可能であることも簡単に述べていた．本研究では，和声学の諸規則を連続関数として再定義することにより，従来の研究では得ることのできなかった，同等に優れたような楽曲に対する，細かな優劣決定の方法を提示した．ここで再定義したこれら和声学の関数の妥当性（それが実際の作曲者の認識と合致するような値や優劣決定を返すかどうか等）については，今後検証が必要である．しかし，和声学知識を客観的に実装した統一的な評価指標として，今後も楽曲の評価にそのまま用いることができるだろうと著者は考えている．他のシステムが生成した楽曲にもこれら関数を適用することで，統一され，また企画された指標を様々な規則群に対して得ることが可能であり，先行研究に基づく新たな研究の発祥と発展に貢献すると考えている．

第5章　おわりに

本章では，本研究の概要と，得られた結果とをまとめる．また，今後の展望についても説明し，これからの具体的な研究の方針も考える．

5.1　まとめ

本稿の貢献は大きく 3 つの部分にまとめることができる．

1 つ目は，多くの分野で孤立的に研究がなされてきた自動作曲研究を多面的に取り上げ，その研究領域の現状を複合的に考察したところにある．特に本稿では，自動作曲研究に関する音楽情報処理分野，音楽学分野，生体心理学分野側面の多くの先行研究をまとめ，その知見を一つに集約し，実用的な自動作曲アプリケーションを構築するにあたって考慮しなければならない様々な課題をまとめた．また，それらの課題に対する具体的なアプローチとして，以下に説明する，ヒューリスティック・ルールベースを用いた初期解の生成，ダイナミック・プログラミングの利用，及び和声学規則の連続関数としての再定義といった，複合的な作曲アルゴリズムを提案した．

2 つ目の貢献として，本稿では，提案した自動作曲のアプローチを実際に自動伴奏生成システムとして実装し，それが先に上げた課題の解決につながることを，具体的な例や客観的な統計解析を用いて説明した．特に，自動作曲の大きな課題である，膨大な探索空間を効率的に扱う手法として，ヒューリスティック・ルールベースを用いた確率モデルによる初期解の生成と，ダイナミック・プログラミングによる解探索といった 2 つのアルゴリズムの併用を用いた，新たな探索手法を提案した．また，提案した手法は，自動作曲システム*”Auto Chorus Creator(ACC)”* として実装した．

3 つ目として，和声学規則を連続関数として再定義したことが挙げられる．一般的な音楽学において，和声に関する諸規則は，破るか破らないかのみを判定する，二値関数の形で定義されている．著者の知る限り，和声学の知識を実装した先行研究でも，これら規則は二値関数のまま実装されている．本稿では，作曲の過程にける許容解の評価に，より明確で客観的な優劣決定を行うために,『芸大和声』に記載されている 13 の規則を連続関数として再定義し，システムの評価関数として実装した．これにより，従来の評価手法では同等に優れている（劣っている）と評価されるような複数の楽曲にも，厳密には異なっているその和声的特徴等を反映する，統一的な評価指標を提案できた．

5.2　今後の展望

本稿では，和声学の諸規則を連続関数として再定義した．しかし，これが実際の作曲における作曲者の思考や，和声学における専門家たちの経験的感覚と合致しているかは検証

していない．和声学規則の違反の度合いや，また違反が許容される状況というのは体系的にまとめられておらず，作曲に携わる各個人が既存の楽曲を解析して，その現象をそれぞれ独自にルール化していくものと考えられる．この過程は，機械学習の手法などを利用することにより，直感的に模倣できるだろうと考えられる．今後は，本稿で定義していない他の規則も連続関数として再定義するとともに，それら評価関数が，専門家による一般的な和声理解をどの程度反映したものであるか調べることが必要である．

また，本稿では，自動作曲における一般的な解決定手法として，2 つのアルゴリズムの併用による複合的なアルゴリズムを提案している．今後は，本研究で扱ったように，古典的和声学が適用される比較的狭い音楽ジャンルの制限を解除し，より広い分野におけるこの手法の有用性を検討したい．その際，初期解の生成や探索評価に利用されるルールベースに，非和声音，リズム，コード置換等，現在実装している音楽学知識の応用に近い楽曲特徴からルールを加えることで，徐々にシステムの作曲の幅を広げ，実際の自動作曲アプリケーションを構築しながらこの手法の可能性を探りたい．

理論上，本システムは 4 声以上の楽曲編成，及び声楽とは異なる楽器編成に同じく拡張することができる．しかしその実施と実際の検証は行っていない．GUI を開発し，これら拡張機能を供え持つようにシステム実装を拡張し，近いうちにシステムを広く聴衆に公開したいと考えている．著者が研究の根底に持っている，「作曲に関する専門知識を持たないユーザの作曲活動を支援したい」という興味を踏まえ，対象とするユーザ層が自動作曲システムに具体的に何を求めるのか等を，ユーザ実験等を通して考察する必要もあるだろう．

最後になるが，一般に自動作曲システムの客観的評価は難しく，統一的な評価指針を示した先行研究も存在しない．本稿を記載するにあたり，著者もいくつかの評価方法を検討したが，探索空間の全探索にかかる時間の膨大さ，作曲アルゴリズムの複雑さ故の性能の理論的解析の困難さ，最適解が未知である故の複数解の客観的比較の困難さ等，様々な障壁に行き当たった．今後も，自動作曲研究の体系的な発展のためにも，自動作曲システムの客観的な評価指針が必要となる．特に，特色ある各研究の様々なシステムを比較するにあたって，ある特定のシステムに特化されない，客観的かつロバストな評価基準が求められている．今後は，旋律の和声要素に限らず，リズムや旋律そのものの評価に用いることのできる評価指針も検討したい．

5.3 謝辞

論文執筆，及び研究活動に際して多大なご指導をいただいた小野哲雄教授，棟方渚助教をはじめとする，研究室の皆さんの協力に，この場を借りて感謝します．

また，2014 年初旬に実習生として迎え入れ，本稿で紹介した自動伴奏生成システムの提案や設計にあたり，ご指導，ご協力をくださった，産業技術総合研究所（AIST）の後藤真孝主席研究員，深山覚研究員，及びメディアインタラクション研究グループの皆さんに感謝します．

更に，研究を始めた当初に音楽情報処理の研究に私を紹介してくださった，北陸先端科学技術大学院大学（JAIST）の東条敏教授，及び，和声学の学習と研究において音楽家として貴重な意見をくださった，ローズガーデンクライスト教会の大野あゆみ先生の助言にも感謝します．また，実験の検討に加わり，作曲者視点からの貴重な意見をくださった，

北海道教育大学の宮下和也氏にもお礼の言葉を述べます.

最後に，日々私を支えてくださっている神と，大学生活 7 年間を見守ってくださった両親と兄弟姉妹にも感謝の言葉を述べます． Thank you for the love and support you have shown me, especially over the past few years of my life. I could not have made it this far were it not for your love and prayers.

参考文献

[1] Hermann Hild and Johannes Feulner Wolfram Menzel. HARMONET: A Neural Net for Harmonizing Chorales in the Style of J.S.Bach. In *Proceedings of Advances in Neural Information Processing Systems 4*, pp. 267–274, 1991.

[2] François Pachet and Pierre Roy. Musical Harmonization with Constraints: A Survey. *Constraints Journal*, Vol. 6, No. 1, pp. 7–19, 2001.

[3] François Pachet. The Continuator: Musical Interaction With Style. *Journal of New Music Research*, Vol. 32, No. 3, pp. 333–341, 2003.

[4] Satoru Fukayama, Kei Nakatsuma, Shinji Sako, Yuichiro Yonebayashi, Tae Hun Kim, Si Wei Qin, Takuho Nakano, Takuya Nishimoto, and Shigeki Sagayama. Orpheus: Automatic Composition System Considering Prosody of Japanese Lyrics. In *Proceedings of International Conference on Entertainment Computing*, pp. 309–310, 2009. Paris.

[5] David Cope. *Machine Models of Music*. MIT Press, 1992.

[6] François Pachet, Pierre Roy, and Gabriele Barbieri. Finite-Length Markov Processes with Constraints. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pp. 635–642, 2011. Barcelona.

[7] Somnuk Phon-Amnuaisuk and Geraint A. Wiggins. The Four-Part Harmonisation Problem: A comparison between Genetic Algorithms and a Rule-Based System. In *Proceedings of AISB'99 Symposium on Musical Creativity*, 1999. Edinburgh.

[8] Kemal Ebcioğlu. An Expert System For Chorale Harmonization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 784–788, 1986. Philadelphia.

[9] John A. Biles. GenJam: A Genetic Algorithm for Generating Jazz Solos. In *Proceedings of International Computer Music Conference*, pp. 131–137, 1994. San Francisco.

[10] Robert M. MacCallum, Matthias Mauch, Austin Burt, and Armand M. Leroi. Evolution of music by public choice. In *Proceedings of the National Academy of Sciences USA*, No. 30 in 109, 2012.

[11] Matthew E. P. Davies, Philippe Hamel, Kazuyoshi Yoshii, and Masataka Goto. AUTOMASHUPPER: An Automatic Multi-Song Mashup System. In *Proceedings*

of the 14th International Society of Music Information Retrieval Conference (ISMIR 2013), pp. 575–580, 2013. Curitiba.

[12] Eric J. Humphrey, Douglas Turnbull, and Tom Collins. Creativity in Music Informatics - Discussion Review. In *the 14th International Society of Music Information Retrieval Conference (ISMIR 2013) late-breaking news and demos*, 2013. Curitiba.

[13] Moray Allan and Christopher K. I. Williams. Harmonising Chorales by Probabilistic Inference. In *Proceedings of the Neural Information Processing Systems Conference*, pp. 25–32, 2004. Vancouver and Whistler.

[14] Syunpei Suzuki annd Tetsuro Kitahara. Four-part Harmonization Using Probabilistic Models: Comparison of Models With and Without Chord Nodes. In *Proceedings of the Sound and Music Conference (SMC 2013)*, pp. 628–633, 2013. Stockholm.

[15] 松原正樹, 深山覚, 奥村健太, 寺村佳子, 大村英史, 橋田光代, 北原鉄朗. 創作過程の分類に基づく自動音楽生成研究のサーベイ. コンピュータソフトウェア, Vol. 30, No. 1, pp. 101–118, 2013.

[16] 池内友次郎, 島岡譲ら. 和声 理論と実習, I,II,III, 別巻. 音楽之友社, 1964.

[17] 江村伯夫, 三浦雅展. 情報技術に基づく作編曲の現状. 日本音響学会誌, Vol. 67, No. 6, pp. 233–238, 2011.

[18] 千蔵八郎. 音楽史＜作曲家とその作品＞, 1:古代からグレゴリオ聖歌の成立まで, pp. 6–8. 教育芸術社, 1983. ISBN:4-905700-38-8 C7073.

[19] Kimberly Smith. *Music and Morals. Dispelling the Myth That Music Is Amoral*, chapter 9:The Decline of Music Through the Ages, pp. 77–79. Winepress Publishing, 2005. ISBN:1-57921-765-6.

[20] 新改訳聖書刊行会（編）. 新改訳 BI-20 中型聖書 ―引照・注付―, 歴代誌第二 5 章 12-13 節, p. 739. いのちのことば社, 2003. ISBN：978-4-264-02001-1.

[21] オリヴィエ・アラン (永富正之, 二宮 正之共訳). 和声の歴史, 2:音程から和音へ, pp. 21–30. 白水社, 1969. ISBN:4-560-05448-7.

[22] 千蔵八郎. 音楽史＜作曲家とその作品＞, 3:中世の音楽 (13 世紀まで), pp. 15–17. 教育芸術社, 1983. ISBN:4-905700-38-8 C7073.

[23] 三浦雅展, 下石橋徹, 斉木由美, 柳田益造. 和声学におけるバス課題についての回答確認システムの構築とその評価. 電子情報通信学会論文誌, Vol. J85-D-II, No. 6, pp. 936–945, 2001.

[24] 三浦雅展, 柳田益造. ソプラノ課題の全許容解列挙システムの構築. 日本音響学会誌, Vol. 60, No. 3, pp. 105–114, 2004.

[25] 三浦雅展, 山田真司, 柳田益造. 四声体和声の音楽美を評価するシステム”MAESTRO”. 日本音響学会誌, Vol. 59, No. 3, pp. 131–140, 2003.

[26] 三浦雅展, 江村伯夫. 和声理論に基づいた作・編曲システム. システム/制御/情報, Vol. 56, No. 5, pp. 213–218, 2012.

[27] Masatochi Hamanaka, Keiji Hirata, and Satoshi Tojo. Implementing “A Generative Theory of Tonal Music”. *Journal of New Music Research*, Vol. 35, No. 4, pp. 249–277, 2006.

[28] 坂本鐘期, 東条敏. Tonal Pitch Space を用いた楽曲の和声解析. 情報処理学会研究報告, No. 9 in 2009-MUS-80, pp. 1–6, May 2009.

[29] Karl F. MacDorman, Stuart Ough, and Chin-Chang Ho. Automatic Emotion Prediction of Song Excerpts: Index Construction, Algorithm Design, and Empirical Comparison. *Journal of New Music Research*, Vol. 36, No. 1, pp. 281–299, 2007.

[30] Gregory D. Webster and Catherine G. Weir. Emotional Responses to Music: Interactive Effects of Mode, Texture, and Tempo. *Motivation and Emotion*, Vol. 29, No. 1, pp. 19–39, 2005.

[31] Marcel Zentner, Didier Grandjean, and Klaus R. Scherer. Emotions Evoked by the Sound of Music: Characterization, Classification, and Measurement. *Emotion*, Vol. 8, No. 4, pp. 494–521, 2008.

[32] 大出訓史, 今出篤, 安藤彰男, 谷口高士. 音楽聴取における “感動” の評価要因—感動の種類と音楽の感情価の関係. 情報処理学会論文誌, Vol. 50, No. 3, pp. 1111–1121, March 2009.

[33] 中村均. 音楽の情動的性格の評定と音楽によって生じる情動の評定の関係. *The Japanese Journal of Psychology*, Vol. 54, No. 1, pp. 54–57, 1983.

[34] Nikki S. Rickard. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, Vol. 32, No. 4, pp. 371–388, October 2004.

[35] Karl F. MacDorman, Stuart Ough, and Chin-Chang Ho. Automatic Emotion Prediction of Song Excerpts: Index Construction, Algorithm Design, and Empirical Comparison. *Journal of New Music Research*, Vol. 36, No. 4, pp. 281–299, 2007.

[36] Konstantinos Trohidis, Grigorios Tsoumakas, George Kalliris, and Ioannls Vlahavas. Multi-Label Classification of Music into Emotions. In *Proceedings of the Ninth International Society for MusicInformation Retrieval Conference (ISMIR 2008)*, pp. 325–330, 2008.

[37] William Forde Thompson, E. Glenn Schellenberg, and Gabriela Husain. Arousal, Mood, and The Mozart Effect. *Psychological Science*, Vol. 12, No. 3, pp. 248–251, 2001.

[38] Filippo Bonini Baraldi, Giovanni De Poli, and Antonio Rodà. Communicating Expressive Intentions with a Single Piano Note. *Journal of New Music Research*, Vol. 35, No. 3, pp. 197–210, 2006.

[39] Rob Weale. Discovering How Accessible Electroacoustic Music Can Be: the Intention/Reception project. *Organised Sound*, Vol. 11, No. 2, pp. 189–200, 2006.

[40] Giovanni De Poli, Antonio Rodà, and Alvise Vidolin. Note-by-Note Analysis of the Influence of Expressive Intentions and Musical Structure in Violin Performance. *Journal of New Music Research*, Vol. 27, No. 3, pp. 293–321, 1998.

[41] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一. RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース. 情報処理学会論文誌, Vol. 45, No. 3, pp. 728–738, 2004.

[42] Gordon C. Bruner. Music, Mood, and Marketing. *Journal of Marketing*, Vol. 54, No. 4, pp. 94–104, October 1990. ProQuest Research Library.

[43] Satoru Fukayama, Daisuke Saito, and Shigeki Sagayama. Assistance for Novice Users on Creating Songs from Japanese Lyrics. In *Proceedings of the International Computer Music Conference (ICMC 2012)*, pp. 441–446, 2012. Ljubljana.

[44] Richard Bellman. On The Theory Of Dynamic Programming. *Proceedings of the National Academy of Sciences USA*, Vol. 38, No. 8, pp. 716–719, 1952.

発表論文

[45] Benjamin Luke Evans, Nagisa Munekata, and Tetsuo Ono. Does the Audience Hear My Heart?–Comparing the Physiological Responses of Listeners with Those of the Composer–. In *Proceedings of the International Conference on Physiological Computing Systems (PhyCS 2014)*, pp. 161–166, January 2014.

[46] ベンジャミン ルカエバンズ, 渚棟方, 哲雄小野. Does the Audience Hear My Heart?–作曲者意図とリスナーの特性による楽曲印象の比較–. 研究報告音楽情報科学 (MUS), No. 4 in 2013-MUS-100, pp. 1–6, August 2013.

[47] ベンジャミン ルカエバンズ, 渚棟方, 哲雄小野. 作曲者意図とリスナーの特性による楽曲印象の比較. 第 32 回情報処理学会エンターテイメントコンピューティング研究会研究発表会, June 2014. session–EC3.

[48] Benjamin Evans, Satoru Fukayama, Masataka Goto, Nagisa Munekata, and Tetsuo Ono. AutoChorusCreator : Four-Part Chorus Generator with Musical Feature Control, Using Search Spaces Constructed from Rules of Music Theory. In *Proceedings of the joint Conference of ICMC and SMC*, pp. 1016–1023, September 2014.

[49] Benjamin Evans, 敏東条, 渚棟方, 哲雄小野. Composing Music for You : ユーザの嗜好を取り入れた和声学に基づく合唱譜自動作成システム. 情報処理学会第 75 回全国大会講演論文集, pp. 2–301–2–302, March 2013.

[50] ベンジャミン ルカエバンズ, 渚棟方, 哲雄小野. 和声学に基づく合唱譜自動作成システム ’ CMY ’ –転回系の実装と評価–. 研究報告音楽情報科学 (MUS), No. 37 in 2013-MUS-100, pp. 1–4, September 2013.

[51] Benjamin Evans, 渚棟方, 哲雄小野. 和声学に基づく合唱譜自動作成システム. 研究報告音楽情報科学 (MUS), No. 13 in 2013-MUS-100, August 2013. ポスター発表.

[52] ベンジャミン ルカエバンズ, 渚棟方, 哲雄小野. 和声学に基づく自動作曲システム ’ CMY ’. 人工知能学会第 28 回全国大会講演論文集, 1K5-OS-07b-3, pp. 1–2, May 2014.

[53] ベンジャミン ルカエバンズ, 渚棟方, 哲雄小野. 自動和声付与システム ’ CMY ’. 情報処理学会北海道シンポジウム, October 2014. session–A-03.

図 目 次

2.1 四声体和声における各声部の基本的な音域 7
2.2 バス課題とその解答の例（『芸大和声』第 I 巻 p.41 課題 9–1） 11
2.3 和声付けの流れの例 11
2.4 自動和声付与システム *CMY* の許容解決定の流れ 13
2.5 楽曲 A,B,C に対する作曲者の評価と，被験者（32 人）の評価平均（エラーバーは標準偏差を表す） 19
2.6 楽曲 A,B に対する被験者（32 人）の評価平均（エラーバーは標準偏差を表す） 19
2.7 楽曲 A に対する作曲者評価と，楽曲 A,C に対する被験者（12 人）の評価平均 20

3.1 *Auto Chorus Creator (ACC)* システム概要図 26
3.2 リード譜の例 （『If You're Happy and You Know It』の冒頭 3 小節分） . . 26
3.3 ヒューリスティック・ルールの例 1 28
3.4 ヒューリスティック・ルールの例 2 28
3.5 自動作曲システム *ACC* のユーザ視点での概要図 35

4.1 自動作曲システム *ACC* のモジュール 1 のみ利用した際の出力例（それぞれ，上からアルト，テナー，バスのパートの出力例を表わす） 38
4.2 自動作曲システム *ACC* のモジュール 2 を，複数の初期解に利用した際の出力例（それぞれ，上からアルト，テナー，バスのパートの出力例を表わす） 40
4.3 モジュール 1（ヒューリスティック・ルールベース）の利用の有無による初期解評価値，及び出力評価値の比較（同じ入力に対する 100 回のシステム実行の平均．エラーバーは標準偏差を示す.） 42
4.4 ヒューリスティック・ルールベースの利用の有無による，複数楽曲に対する最終的な出力の解評価値の比較（同じ入力に対する 100 回のシステム実行の平均．エラーバーは標準偏差を示す.） 44

表 目 次

2.1 バスパート譜を入力として与えた際に，和声付けシステムが各種条件下で生成し得る楽曲の理論数（n は入力のバス音の数を示す） 8
2.2 Bruner の 9 つの感情因子 [42] と，本研究で用いた訳語，及び各感情を誘起する音楽要素とそれらの特徴 17

3.1 *ACC* 第 2 モジュールの評価関数に用いた各変数の説明 36

4.1 モジュール 1 に生成された初期解（図 4.1），及びモジュール 2 に生成された出力（図 4.2）に対する，各項目の評価値の比較 41