

IRMに基づくジャズアドリブの構造分析
-バップとモード間の比較-

筑波大学
図書館情報メディア研究科
2020年3月
池田 周平

目次

第1章	序章	1
1.1	ジャズ, バップ, モードジャズ	1
1.2	論文構成	1
第2章	音楽理論	2
2.1	バップ, モードに関連する音楽理論	2
2.1.1	機能和声	2
2.1.2	モード理論	3
2.2	Implication Realization Model	5
第3章	関連研究	7
3.1	音楽生成モデル	7
3.2	ジャズの音楽的理解	8
3.3	本研究の位置付け	8
第4章	手法	10
4.1	分析データ	10
4.1.1	データ形式	11
4.2	セグメントへの分割手法	12
4.3	確率的文脈自由文法	12
4.3.1	文法の構築	13
4.3.2	パラメータ推定	15
第5章	結果と考察	17
5.1	サンプリング結果	17
5.1.1	頻度分析	17
5.1.2	IRM シンボル別分析	21
5.1.3	各非終端記号における導出規則選択確率	24
5.2	文法から導出不可能なセグメント	25
5.2.1	バップの導出不可能セグメント	25
5.2.2	モードの導出不可能セグメント	26
第6章	終章	28
6.1	まとめ	28
6.2	今後	29
	参考文献	31
付録A	分析に使用した曲	35

付録 B 分析時のパラメータの詳細	37
B.1 セグメンテーション時のパラメータ	37
B.2 学習時のパラメータ	37
付録 C 学習結果	39

目 次

2.1	ダイアトニックコード (C が主音の場合)	2
2.2	Miles Davis, Donna Lee	2
2.3	5 度圏	3
2.4	教会旋法	4
2.5	D ドリアンモード	4
2.6	Miles Davis, So What	4
2.7	各シンボルの音形	5
2.8	IRM の分析例 1(Bach, Fugue 16, WellTempered Clavier, 2, mm. 1-4.)	6
2.9	IRM の分析例 2(Bach, Fugue 2, WellTempered Clavier, 2, m. 1.)	6
4.1	確率モデル分析の流れ	10
4.2	ジャズ楽曲の進行例	11
4.3	譜例	11
4.4	譜例	13
4.5	音程系列導出の例	14
4.6	非終端記号導出の様子	15
5.1	バップにおける導出規則上位 60	17
5.2	モードにおける導出規則上位 60	18
5.3	バップの音程集計	18
5.4	モードの音程集計	18
5.5	非終端記号の出現回数	19
5.6	$\lambda_{I[n]}$	19
5.7	バップの特徴的導出規則の考察	20
5.8	バップの P 系導出規則上位 30	21
5.9	モードの P 系導出規則上位 30	21
5.10	バップの R 系導出規則上位 30	21
5.11	モードの R 系導出規則上位 30	21
5.12	Charlie Parker, Embraceable You m.19	22
5.13	Charlie Parker, Don't Blame Me m.1	22
5.14	減七和音とそれに基づく音階	22
5.15	Bob Berg, Nature Beast, m.63	23
5.16	Kenny Wheeler, Pass It On, m.16	23
5.17	$\theta_{I[-7]}$	24
5.18	$\theta_{I[7]}$	24
5.19	$\theta_{I[-12]}$	24
5.20	$\theta_{I[12]}$	24
5.21	Charlie Parker, Don't Blame Me, m. 11.	25

5.22	Charlie Parker, EmbraceableYou, mm. 22–23	25
5.23	Bob Berg, Nature of the Beast, mm. 87–88.	26
5.24	Cannonball Adderley, So What, mm. 49–51.	26
5.25	Chris Potter, Arjuna, m. 99.	26
5.26	George Coleman, MaidenVoyage, mm.13–15.	27
5.27	Chris Potter, PopTune, m. 50.	27
5.28	Chris Potter, PopTune, mm.110–111.	27
B.1	暗意の強さの図式 (文献 [7]Fig.5.1 を参考に作成)	38
C.1	$\theta_{I[-16]}$	39
C.2	$\theta_{I[0]}$	39
C.3	$\theta_{I[-15]}$	39
C.4	$\theta_{I[15]}$	39
C.5	$\theta_{I[-14]}$	40
C.6	$\theta_{I[14]}$	40
C.7	$\theta_{I[-13]}$	40
C.8	$\theta_{I[13]}$	40
C.9	$\theta_{I[-12]}$	40
C.10	$\theta_{I[12]}$	40
C.11	$\theta_{I[-11]}$	40
C.12	$\theta_{I[11]}$	40
C.13	$\theta_{I[-10]}$	41
C.14	$\theta_{I[10]}$	41
C.15	$\theta_{I[-9]}$	41
C.16	$\theta_{I[9]}$	41
C.17	$\theta_{I[-8]}$	41
C.18	$\theta_{I[8]}$	41
C.19	$\theta_{I[-7]}$	41
C.20	$\theta_{I[7]}$	41
C.21	$\theta_{I[1-6]}$	42
C.22	$\theta_{I[6]}$	42
C.23	$\theta_{I[-5]}$	42
C.24	$\theta_{I[5]}$	42
C.25	$\theta_{I[-4]}$	42
C.26	$\theta_{I[4]}$	42
C.27	$\theta_{I[-3]}$	42
C.28	$\theta_{I[3]}$	42
C.29	$\theta_{I[-2]}$	43
C.30	$\theta_{I[2]}$	43
C.31	$\theta_{I[-1]}$	43
C.32	$\theta_{I[1]}$	43

第1章 序章

1.1 ジャズ, バップ, モードジャズ

ジャズとは、19世紀末から20世紀初頭のアメリカ南部で形成された音楽をルーツとした音楽文化や様式である [1][2]。その中でも、1950年以降に発展した即興演奏が積極的に行われる演奏様式を指してモダンジャズと呼称する [3]。モダンジャズでは、モダン以前のジャズに比べて音楽理論的な専門性や旋律の複雑さが大きくなっている [4]。モダンジャズを研究することには、以下の意義が存在する。

- 即興では、人間の音楽認知のプロセスと創造のプロセスが同時に進行している。モダンジャズの発展期に残された演奏を分析することで、音楽における認知と創造の関連性解明につながる。
- モダンジャズは近代に発展した。ゆえに伝記や録音が詳細に残っており、歴史的背景や音楽理論的な分析を裏付けとした詳細な研究ができる。

本論文では以降、ジャズという言葉は主にモダンジャズの様式に則った音楽を示すために用いる。

ジャズの発展に伴い、現在では様式の細分化がされている。本研究では、ジャズの内のバップと呼ばれる様式とモードと呼ばれる様式を対象としている。以下にその概要を述べる。

バップ 1950年代前半、ジャズプレイヤー間で、即興演奏を複雑・細分化させる動きが存在した。そのような、Charlie Parker や Dizzy Gillespie らに代表される機能和声の理論に基づくジャズのあり方をバップと呼称する¹。

モード 1950年代終盤から60年代にかけて、バップから抜け出そうとする動きが生じた。そのような演奏方式は、Miles Davis のアルバム「Kind of Blue」によって明示的に打ち出された [4]。このアルバムから発展した、モードと呼ばれる音組織に基づくジャズの様式をモードジャズ、または単にモードと呼称する。

バップとモードは、即興演奏の方針だけでなく、基礎となる音楽理論が異なっていることにも注目される。本研究では、この2種類のジャズの様式で演奏された記録を分析することで、ジャズにおける認知と創造過程が音楽理論とどう関わっているかを分析する。

1.2 論文構成

本論文の論文構成を述べる、2章では本研究の基本となる音楽理論に関する基本事項を述べる。つづいて3章にて、関連研究の事例を挙げる。具体的には、音楽に対する記号処理モデル作成する手法を用いた研究事例と、ジャズの音楽的理解を目的とした研究事例を挙げている。その後、本研究の立ち位置についても触れる。4章では、本研究で用いたモデルの構築手法について述べる。5章で分析結果とその考察を述べた後、6章でまとめと今後の展望を述べる。

¹Be bop, ビバップという呼称も存在するが、本論文ではバップと統一する

第2章 音楽理論

2.1 バップ、モードに関連する音楽理論

本節では、バップとモードの背景に存在する音楽理論について述べる。2.1.1 節で、バップの基礎となる機能和声を、2.1.2 節でモードの基礎となるモード理論を記述する。

2.1.1 機能和声

バップにおける即興演奏では機能和声の考え方が積極的に取り入れられている。

機能和声とは、ダイアトニックコードと呼ばれる7種の和音を基本とする音楽理論の体系である。ダイアトニックコードの例を図2.1に示す。ダイアトニックコードを構成するには、和声の中心となる音(主音)が必要となる。図2.1の場合は、主音をCとしている。それぞれの和音にはTonic(安定), Dominant(緊張), Subdominant(弱い緊張)の3種類の役割が割り振られている。それらの役割や、和音と和音の代理関係(例えば、同じ役割の和音は交換可能など)のルールを組み合わせて作曲やアドリブを行う。

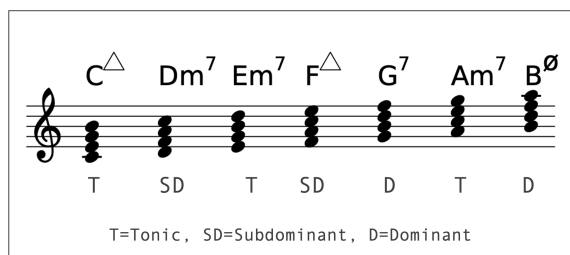


図 2.1: ダイアトニックコード (Cが主音の場合)

バップにおける機能和声の活用例として、Miles Davisがモードに移行する以前の1947年に録音したDonna Leeの終盤を譜例2.2に挙げる。

図 2.2: Miles Davis, Donna Lee

楽譜の上段とアルファベットは伴奏に使用される和音を示す。下段の旋律に使用されている音と、伴奏の和音に含まれる音が多く共通している。また、図の赤枠で囲んだ箇所は、和音が変化するタイミングで前の和音の7度の音から次の和音の3度に移行する

動きになっている。これによって、伴奏で指定された和音の進行 (Cm7→F7→Bbm7) を旋律で感じさせることができる。

譜例には和音の進行にもバップによく見られる特徴が存在する。和音を構成する中での最低音の動きは C → F → Bb → Eb → Ab のように、図 2.3 に示す 5 度圏を反時計回りに遷移している。この時、以下のように和音の働きを解釈できる。

- C → F → Bb は Bb を主音とした Subdominant → Dominant → Tonic の動きである (主音からの距離をローマ数字で表し、Bb を主音とした II → V → I とも表現される)
- F → Bb → Eb は Eb を主音とした Subdominant → Dominant → Tonic の動きである (Eb を主音とした II → V → I)
- Bb → Eb → Ab は Ab を主音とした Subdominant → Dominant → Tonic の動きである (Ab を主音とした II → V → I)

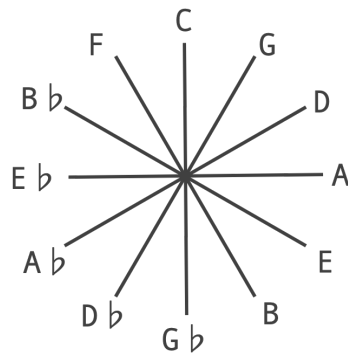


図 2.3: 5 度圏

和音の働きをまとめると、局所的な主音を頻繁に変更しながら II → V → I と表現される進行が連鎖していることがわかる。このような和声もバップの即興演奏において重要となる。

2.1.2 モード理論

モードとは旋法と訳され、音階組織やそれに基づく旋律のあり方を指す [5]。教会旋法 (図 2.4) や Lydian Chromatic Concept [6] のように、機能 and 声のシステムとは独立とされる。

図 2.4 は 7 種類の教会旋法を、軸音 (モードの核となる音、終止感を感じるとされる。) を C として表示した図である。モードは、軸音との音程関係によって構成される。

例えば、ドリアンモードは軸音に対して 1 度、長 2 度、短 3 度、完全 4 度、完全 5 度、長 6 度、短 7 度の 7 音で構成される。この音程関係から、ドリアンモードの特徴 (大きくは以下に示す 2 点) が決定される。ドリアンモードに基づく即興演奏や作曲では以下の 2 点が重要視される。

1. 先頭から 3 番目の音が短 3 度である。これによって、機能 and 声における短調のような響きとなる。
2. 先頭から 6 番目の音が長 6 度である。対して他の短 3 度を持つ教会旋法フリギアン、エオリアン、ロクリアンは短 6 度である。この差異によって、6 度の音がドリアンモードを特徴づける音とされる。



図 2.4: 教会旋法

ドリアンモードの音程関係をCの音に対して適用したのが図の2段目である。同じ音程関係をDに適用することで、Dドリアンモードという新しいモードを構成することができる。



図 2.5: Dドリアンモード

モードジャズにおいては、和音ではなくモードの構成音を元に、楽曲やアドリブを演奏する。例として、Miles Davis 作の So What の冒頭¹をに挙げる。



図 2.6: Miles Davis, So What

この曲はDドリアンモードを主に書かれている。そのため、BやCの音が臨時記号 \sharp , \flat なしで用いられている。機能和声の基本に従っていた場合は、伴奏の和音Dmの働きを強調する、または邪魔しない音高が選択され、 $B\flat$ や $C\sharp$ となる。

また、この楽曲は、譜例 2.6 を繰り返しながら展開していく。つまり、楽曲の大部分が同じモードから変わらないことを意味する。この点も、局所的な音楽の重心が頻繁に転じるバップとは対照的である。

¹説明のため簡約化してある

2.2 Implication Realization Model

Implication Realization Model(IRM) は, Narmour によって提唱された認知的音楽理論である [7]. この理論は, 暗意実現モデルと訳され, 文字通り暗意とその実現に着目し旋律を分析する.

暗意とは, 2つの音によって1つの音程が形成された際, 次の音程を無意識に期待することを指す. 暗意に重要な原理として, (1) 音程の広さに関する Intervalic Difference(音程原理) と (2) 音程の向きに関する Registral Direction(方向原理) の2種類が存在する. それぞれ, 狭い音程・広い音程に対し, 2.1 に示すような異なる作用をもつ.

表 2.1: IRM における音程原理と方向原理

聴こえた音程	Intervalic Difference	Registral Direction
狭い音程	次の音程も狭いと暗意させる	同符号の音程を暗意させる
広い音程	次の音程は狭いと暗意させる	異符号の音程を暗意させる

生じた暗意が, 3つ目の音によって実現されるか, 裏切られるかを基に, 表 2.2 に示す 8 種類の音形を基本シンボルとして定義する. それぞれのシンボルに分類される音形の例を図 2.7 に示す.

表 2.2: IRM における基本シンボル

シンボル名	2つの音程の向き	1つ目の音程	2つ目の音程	説明
P	同じ	狭い	狭い	進行
IP	逆	狭い	狭い	方向に変化をつけた進行
VP	同じ	狭い	広い	音程に変化をつけた進行
R	逆	広い	狭い	逆行
IR	同じ	広い	狭い	方向に変化をつけた逆行
VR	逆	広い	広い	音程に変化をつけた逆行
ID	逆	$\pm n$	$\mp n$	音高変化を含む反復
D	-	0	0	反復



図 2.7: 各シンボルの音形

その他例外的に, 2つの音のみで構成されるシンボル Dyad, 単音で構成されるシンボル Monad が存在する.

表 2.2 に示した基本シンボルは, 3つのグループに分類できる. 本研究ではグループの名前を P系, R系, D系とした.

P系 旋律の進行 (Process) を意味するグループ. Pを基本とし, IP, VPはその亜種である. 人間が容易に受容できる音形であり, 旋律の連続性を感じさせるとされる.

R系 旋律の逆行 (Reversal) を意味するグループ. Rを基本とし, IR, VRはその亜種である. 楽曲の盛り上がりや展開時によく見られる [8].

D系 旋律の反復 (Duplication) を意味するグループ. Dを基本とし, IDはその亜種である. 旋律を聴いている時に反復が来ると, その次も反復が続くと暗意される. 反復が長く続けば続くほど暗意は強くなり, 反復が途切れた際の意外性が高くなる性質を持つ.

シンボルは隣接する3つの音だけでなく, 音符で隔たれた3音を選択して構成することも可能である. Narmour 自身による, 階層的な分析例を, 図 2.8, 図 2.9 に示す.

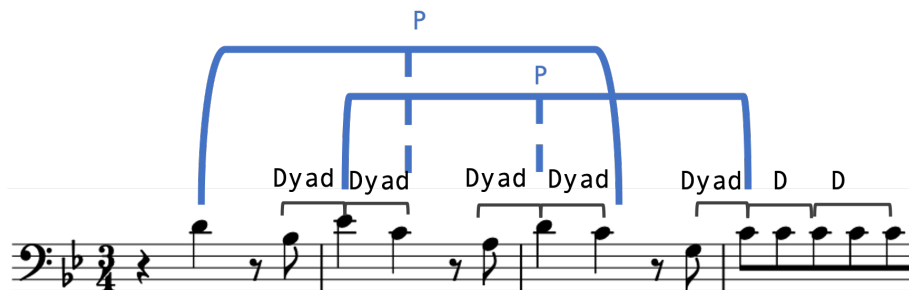


図 2.8: IRM の分析例 1(Bach, Fugue 16, WellTempered Clavier, 2, mm. 1-4.)

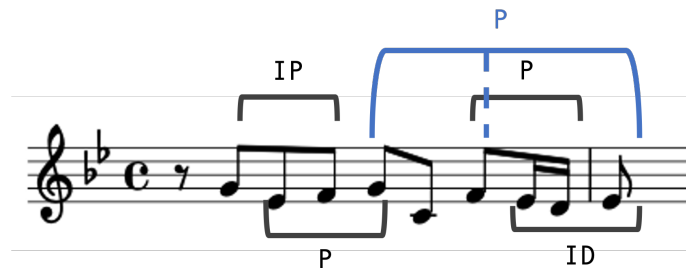


図 2.9: IRM の分析例 2(Bach, Fugue 2, WellTempered Clavier, 2, m. 1.)

図 2.8, 図 2.9 は, 参考文献 [7] の p.133 の図を参考に作成した. 括弧でシンボルを形成する音の範囲を示している. 図 2.9 において上下に配置したシンボルは, 音を共有してシンボルを形成している. 青色で示した括弧のように, 離れた3音を用いてシンボルを構成することも許されている.

第3章 関連研究

3.1 音楽生成モデル

音楽は音響信号として我々に認知されるが、楽譜のような離散記号としても表現可能である。楽譜による作曲や編曲は離散記号系列の操作と言える。現在のデスクトップミュージックにおける音楽制作のような場では、楽譜を用いずに、録音された音を組み合わせる音楽を制作するケースも存在する。このような場合も、録音された音響信号はなんらかの離散記号(音を再生するためのスイッチや音ごとの識別番号)に結びつけられていると解釈できる。つまり、音楽の創造過程や意味構造は、記号表現された音楽から読み取ることができる。記号表現された音楽の解析は、音楽理解において必須であるとも言われている [9]。また、記号上の構造解析を通して、明文化されていない音楽構造を明らかにするという学術的な意義も存在する [10]。

本研究も、記号化された音楽を対象とした研究である。本研究では、「記号化された音楽を生成できるモデル(生成モデル)を構築し、生成された音楽またはモデルの構造から知見を得る」という研究方針をとっている。生成モデルを対象とした研究は広く存在する。以下に生成モデルによく用いられてきた手法と研究例を述べる。

まず、音楽学的な知識や実際の作品をもとに音楽を生成する手法について述べる。そのような方法では、「4小節目にパターンAの旋律が演奏されたら、5小節目にはパターンBを選択する」といった操作を繰り返して楽曲が生成される [11]。与える知識やデータベースによって、自己の演奏を発展させて即興演奏を展開するシステム [12] や、人間の演奏の盛り上がりに合わせて伴奏を出力するシステム [13]、与えた和声をジャズ和声に編曲するシステム [14] のような幅広い応用が可能である。

音楽の高い曖昧性をモデルに組み込む際に、確率を取り入れた手法が用いられている。矢澤らは、旋律をIRMのシンボル列に置き換え、その推移を確率的に表現した [15]。IRMのシンボルを用いることによって、類似した印象を与えるが、細部は異なるような旋律の生成が可能となる。Gillickらは、旋律の抽象化方法を独自に提案し、抽象化された旋律をSlopesと呼称した。Slopesの推移をマルコフモデルで表現し、遷移確率を学習することにより、演奏者の特徴を模倣した即興演奏を生成できる [16]。これら2つの事例は、複数の音が1つの記号に抽象化されている。対して、Sertanらのように、音1つ1つの推移をマルコフモデルとして表現することも可能である [17]。マルコフモデルは系列の前後関係性を記述することができるので、時間軸にそった音楽の展開を表現することができる。

確率的文脈自由文法(PCFG)に基づく音楽生成モデルを利用する手法も盛んである。PCFGは音楽における階層構造やパターンを表現できる。Rohrmeierら [18] や Harasimら [19] は、機能和声に存在する和音の代理関係や進行のきまりを確率モデルで表現し、複雑とされるジャズ和声の解析に成功している。Kellerらは、和声ではなく、音符を単位として確率モデルを構築した。具体的には、「4分音符の長さは、8分音符2つ分の長さと同じ」といった音価の関係を文法とし、モデルに組み込んだ [20]。

音楽の構造分析に用いられるGenerative Theory of Tonal Music(GTTM)が存在する [21]。GTTMでは、人間の音楽認知過程に基づいた音楽の簡約が行われる。ここでいう簡約とは、

旋律の印象や構造に大きく影響する音を選別し、そうでない音を間引きするような操作である。GTTMは旋律の構文解析やモーフィングに応用することができる[22]。しかし、GTTMの簡約のルールは曖昧で、計算機への実装が困難である。この曖昧性の課題に対する解決策としても、PCFGが用いられている[23][24]。

音楽のルールを獲得する法として、遺伝的アルゴリズムを用いる例[25][26][27]も存在する。遺伝的アルゴリズムには、生成された結果を評価するフィッティング関数が必要となる。評価基準には人間の主観[25]や、音楽学によって定められた禁則[27]が用いられている。それらの方法では、バップとモードを同等に評価することは困難である。一方、確率モデルでは評価の必要はなく、音価の関係や音程の推移傾向など、バップとモードに共通する要素をもとにモデルを構築できる。そこで本研究では確率モデルを採用した。

3.2 ジャズの音楽的理解

本節では、ジャズの音楽的理解を深めることを目的とした研究について述べる。

ジャズ即興演奏を対象に分析し、階層的な構造を考察した事例が存在する[28][29]。これらの分析はシェンカー理論に基づいて行われた。シェンカー理論はGTTMと同様、音楽を簡約化して構造分析する理論である。

MartinはCharlie Parkerの即興演奏の音形を分類し、原曲部分と即興演奏部分の関連性を指摘した[30]。SchullerもSonny Rollinsの即興演奏に対し類似の分析を行っている[31]。これらの研究は手動で分析を行った事例であり、客観性・普遍性に欠けるといった批判も存在する。

そこで、計算機を介して普遍的な法則や、未だ明文化されていない法則の発見を目指すようなジャズ演奏理解の試みが行われている。Ramírezらは、表情のない楽譜通りの演奏を、人間のプレイヤーの演奏に近づけるような法則を遺伝的アルゴリズムを用いて学習した[32]。学習した法則を用いて、音の長さや高さからジャズらしい演奏表情をつけることができる。

Belindaは、演奏中のピッチクラス(オクターブを無視した12種類の音高)の分布を、確率混合モデルで表現した[33]。獲得したクラスの混合率や時間的遷移を可視化することで、演奏中の演奏パターンの移り変わりを分析できる。同様に、Snydalの「Improviz」[34]やKellerらの「Impro-Visor」[20]も即興演奏の情報を可視化して分析を支援するシステムといえる。

Norgaardらは、Charlie Parkerの演奏を計算機によって集計し、演奏の冒頭に共通する音程のパターンの存在を指摘した[35]。Jazzomat projectも、音程のパターンを検索、分析できるシステム「Dig the lick」を公開している[36]。これらの研究事例は、即興演奏は階層的なパターンをもとに構造化・分析が可能という考えを支持する。実際に、計算機による即興演奏の動機分析も行われている[37]。

3.3 本研究の位置付け

本研究は、音楽生成モデルを構築し、モデルからジャズの音楽理解を目指す研究として位置している。音楽の階層構造を木として表現する考えに基づき、モデルには確率的文脈自由文法を選択した。文法構築において、展開や音形の意味付けのしやすさからIRMの原理を参考にする。文法獲得のための学習は教師なしでおこなう。

本研究の独自性は、学習をモードとバップに分けて行う点にある。先行研究では機能と声楽を前提に文法や学習データが設定されていた。機能と声楽は西洋を中心に発展・整理された音楽システムである。対するモードは全世界普遍的に存在していると言われる。人間の音楽に対する普遍的な認知・創造過程を考察するためにも、モードを和声と対等に対象とする観点は重要である。また、比較分析を行うことで互いの理論上で新たな知見が得られることを期待する。

第4章 手法

本章では分析に用いた手法の詳細を述べる。まず分析全体の流れを図 4.1 に示す。

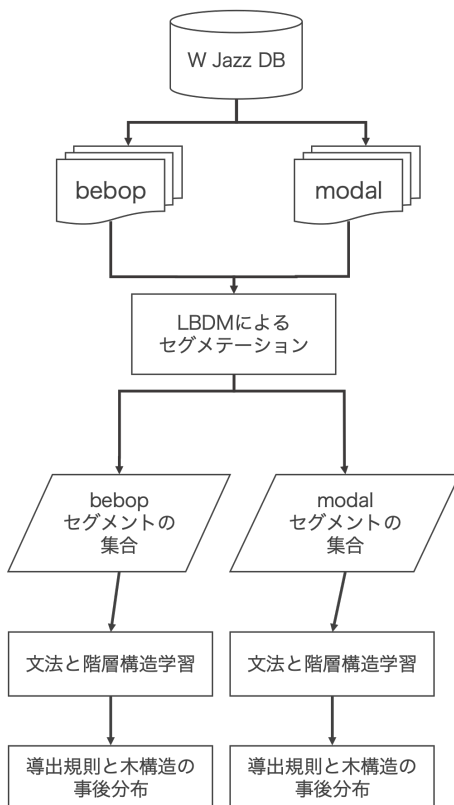


図 4.1: 確率モデル分析の流れ

まずデータベースからバップとモードの即興演奏データを入手し、それぞれ短く分割する。分割された即興演奏の断片をセグメントと呼称する。その後、バップとモードジャズのセグメントを別々に学習し、学習後の文法構造の違いを考察する。

これより、各工程の詳細を述べる。4.1 節ではデータの入手先とデータ形式を、4.2 節ではセグメントへの分割方法を述べる。4.3.1 節にて本研究における文法を定義し、4.3.2 節にて学習法を述べる。

4.1 分析データ

本研究の分析データは、Weimar Jazz Data Base から入手したものを用いる。これは Pfeiderer ら [38][39] が、ジャズの音源中の即興演奏 (アドリブ) を手作業でアノテーションしたデータベースである。Pfeiderer らは、オーディオ分析ソフトを併用し、さらにダブルチェックを行うことで信頼度の高い高解像度データベースを作成できたと主張した。

このデータベースは、ジャズ音源の自動採譜 [40] や、即興演奏支援デバイス [41] の学習コーパスとして成果を挙げている。

4.1.1 データ形式

一般的なジャズの楽曲は図 4.2 のように進行する。

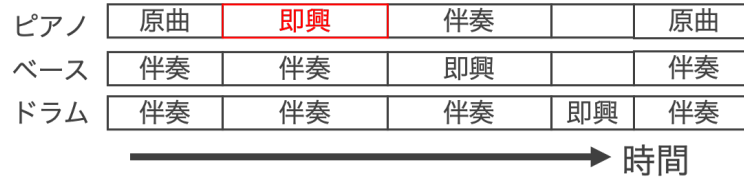


図 4.2: ジャズ楽曲の進行例

図 4.2 の横方向は時間の進みを表す。各楽器が交代で即興演奏を行う。Weimar Jazz DB では、個人のアドリブ (図 4.2 における赤枠で囲まれた範囲) ごとにデータを作成している。

Weimar Jazz DB からは、演奏情報を電子楽器間でやりとりするための規格である MIDI 形式のほか、楽譜描画言語である lilypond 形式、XML 形式、csv 形式でのデータの入手が可能である。本研究では csv 形式のデータを利用する。csv 形式では、1 つの音符が 1 行に対応し、音長や音高など、音符が表現する様々な属性の値が各列に存在する。csv 表現の例を譜例 4.3 を使って示す。



図 4.3: 譜例

譜例 4.3 の 4 分音符は音価に対して 90%，テンポは BPM60 (一拍が一秒) で演奏されたとする。このとき、譜例の csv 表現は表 4.1 のようになる。

表 4.1: csv 表現の例

pitch(音高)	onset(発音時刻)	duration(長さ)
60	0.0	0.9
60	1.0	0.9
67	2.0	0.9
67	3.0	0.5
68	3.5	0.5
69	4.0	0.9
69	5.0	0.9
67	6.0	2.0

4.2 セグメントへの分割手法

Cambouropoulos[42] の提案した Local Boundary Detection Model をセグメント分割手法として用いた。ノート数 $n + 1$ の演奏を想定した分割手法の詳細を以下に述べる。

1. 旋律に含まれる隣接する 2 音の間で、音程, Inter Onset Interval, Offset Onset Interval それぞれについて数列 $P_k = [x_1, x_2, \dots, x_n], k = \{pitch, ioi, ooi\}$ を求める。

音程 2 音の内、後者の音高から前者の音高を引く。単位は半音である。

Inter Onset Interval 後者の音の鳴り始めの時刻から前者の音の鳴り始めの時刻を引く。単位は秒

Offset Onset Interval 後者の音の鳴り始めの時刻から前者の音の時刻を引く。単位は秒

2. $P_{pitch}, P_{ioi}, P_{ooi}$ それぞれにおいて、任意の最大値 $Max_{pitch}, Max_{ioi}, Max_{ooi}$ を超える値は最大値まで切り捨てる。
3. $P_{pitch}, P_{ioi}, P_{ooi}$ の各要素を $Max_{pitch}, Max_{ioi}, Max_{ooi}$ で割ることで $[0, 1]$ の範囲に正規化する。
4. 式 (4.1) にしたがって s_i を求める。
5. それぞれの s_i に重み $w_{pitch}, w_{ioi}, w_{ooi}$ をかけて足し合わせる。
6. 任意の閾値以上の s_i を持つ箇所を演奏を分割する。

s_i の定義を次に示す。

$$s_i = x_i \times (r_{i-1,i} + r_{i,i+1}) \quad (4.1)$$

$r_{i,i+1}$ の定義を次に示す。

$$r_{i,i+1} = \begin{cases} \frac{|x_i - x_{i+1}|}{x_i + x_{i+1}} & (x_i + x_{i+1} \neq 0) \\ 0 & (x_i = x_{i+1} = 0) \end{cases}$$

以上の手法を用いて即興演奏を分割したのち、バップは音符の数が 4 以上 10 未満のものを学習用のセグメントとした。モードは 4 以上 11 未満に制限した。音符の数を制限した理由は、短い音楽的な単位を対象とするの学習を目的とするからである。また、学習時間が短縮できるという理由もある。

最終的にバップは 185 個、モードは 229 個のセグメントを学習に利用する。

4.3 確率的文脈自由文法

本研究のモデルに用いた確率的文脈自由文法 (PCFG) とは、与えられた文法に従う記号列を確率的に生成することのできる確率モデルである [43]。

PCFG の骨格となる文脈自由文法 (CFG) は終端記号の集合 T , 非終端記号の集合 N , 開始記号 S ($S \in N$), 導出規則の集合 R によって構成される。

終端記号とは、生成される記号列を構成する記号である。英語文を対象として CFG を構築した場合、アルファベットや単語が終端記号に相当する。非終端記号とは、終端記号や非終端記号を導出することができる記号である。生成される記号列に非終端記号は存在しない。英語の場合、「名詞」や「名詞句」といった文法上の表現が非終端記号にあたる。

導出とは、「名詞句 ⊃ 形容詞 + 名詞」や「名詞 ⊃ cat」のような関係を利用し、左辺を右辺へと変換することを指す。その導出関係を $A \rightarrow BC$ ($A, B, C \in N$) や $A \rightarrow w$ ($w \in T$) のようにまとめたものが導出規則である。導出規則 R は、ある非終端記号が、どのような終端記号ないし非終端記号を生成するかを記述している。この導出規則 R に、確率的なパラメータ θ を割り当て、確率的に終端記号列を生成できるようにしたモデルが PCFG である。

4.3.1 文法の構築

本研究では音程の系列を導出する文法を構築する。音程の系列とは、旋律の隣り合う音高の差分によって求められる整数列を意味する。

音程系列の例



図 4.4: 譜例

譜例 4.4 に示した MIDI ノートナンバー¹は、左から (60, 60, 67, 67, 68, 69, 69, 67) である。この数列の隣り合う要素の差をとった (0, 7, 0, 1, 1, 0, -2) が譜例から求められる音程系列となる。音程系列に変換することには以下の 2 つの利点が存在する。

1. 時間情報を抽象化し、学習する演奏間に存在する時間依存の差異を考慮する必要がなくなる
2. 相対的な音高変化を見ることで調性を考慮する必要がなくなる

本研究では、Gilbert のモデル [44] のように、音程の値と大文字アルファベットの I を組み合わせた記号を非終端記号とする。親のもつ音程の値が子の値の和となるような関係を導出規則とする。図 4.4 を導出できる木の例を図 4.5 に示す。

数値的な制約のみでは音楽の文法とは言えないので、音楽理論の観点を踏まえた文法を構築する。Gilbert は音符の追加や挿入といった作曲時の操作を基に文法を構築した。対して本研究では、IRM の基本シンボルを基に文法を構築した。文法の定義を以下に示す。開始記号 S は、音程系列の先頭と末尾の和に等しい n を持つ $I[n]$ が選ばれる。

$$N = \{I[n]\} \dots (-16 \leq n \leq 16) \quad (4.2)$$

$$T = \{-16, -15, \dots, 15, 16\} \quad (4.3)$$

$$R = \{P, IP, VP, R, IR, VR, D, ID, repeat, emission\} \quad (4.4)$$

導出規則 R は、以下に示す条件に従って、非終端記号 $I[n]$ ごとに定義される。

¹ピアノの鍵盤の中央のドを 60 として音高を数値化した値。単位は半音

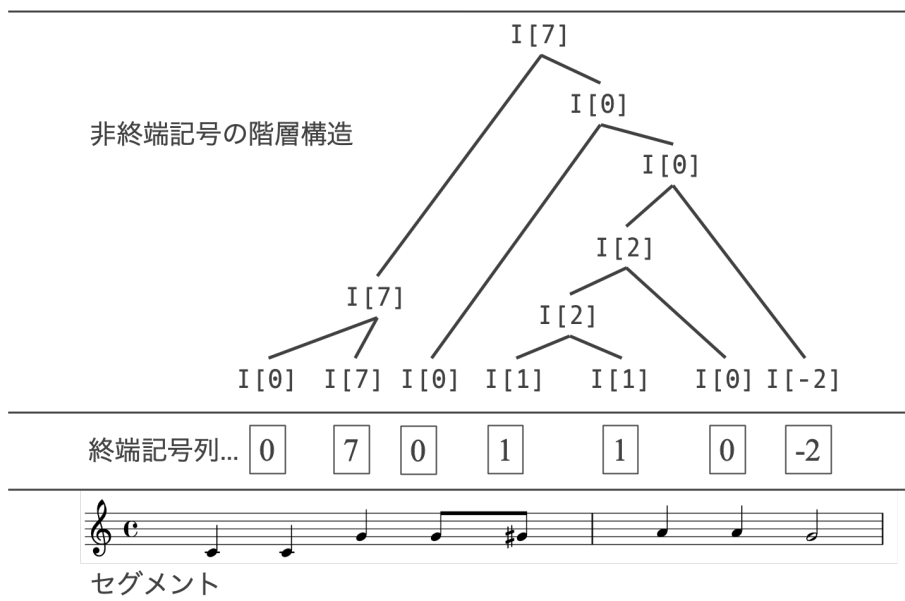


図 4.5: 音程系列導出の例

導出規則の成立条件

$$P : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, |n_1| < 7, |n_2| < 7, n_1n_2 > 0)$$

$$IP : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, |n_1| < 7, |n_2| < 7, n_1n_2 < 0)$$

$$VP : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, |n_1| < 7, 6 < |n_2| < 16, n_1n_2 > 0)$$

上記 3 種類の条件を満たす導出規則から生成される 2 つの音程は、IRM における P 系の音形に相当する。P 系の音形は、未来の音程が、現在聴こえた音程と類似していると暗意され、その暗意が実現した時に生じる。P 系の導出規則は、旋律中の類似した進行を解析することを意図している。

$$R : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, 6 < |n_1| < 16, |n_2| < 7, n_1n_2 < 0)$$

$$IR : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, 6 < |n_1| < 16, |n_2| < 7, n_1n_2 > 0)$$

$$VR : I[n] \rightarrow I[n_1]I[n_2] \dots (n_1 + n_2 = n, 6 < |n_1| < 16, 6 < |n_2| < 16, n_1n_2 < 0)$$

上記 3 種類の条件を満たすルールから導出される音程は、IRM における R 系の音形に相当する。R 系の音形は、未来の音程が現在聴こえた音程と異なっていると暗意され、その暗意が実現した時に生じる。R 系の導出規則は、旋律中の意外性のある進行を解析することを意図している。

$$D : I[0] \rightarrow I[0]I[0]$$

$$ID : I[0] \rightarrow I[n_1]I[n_2] \dots (n_1 = -n_2)$$

上記 2 種類は IRM における D 系のシンボルに相当する音形を導出する。D 系の音形は両端の音高が等しくなる。

以上 IRM の基本シンボル 8 種類に加え, *emission* と, *repeat* を追加する.

$$\begin{aligned} \text{repeat:}I[n] &\rightarrow I[n]I[0]\dots(n \neq 0) \\ &\rightarrow I[0]I[n]\dots(n \neq 0) \\ \text{emission:}I[n] &\rightarrow n \end{aligned}$$

emission は終端記号を出力するルールである. PCFG の定義上必要であったので追加した. *repeat* は同音高を繰り返すためのルールである. IRM の基本シンボルをもとに定義した本研究の文法では導出できない音形を導出する意図に加え, 同音連打を解析する意図が存在する.

非終端記号導出の例として, $I[7]$ が R ルールの $I[7] \rightarrow I[8]I[-1]$ を選択する際の木構造と, 対応する譜面を図 4.6 に示す.

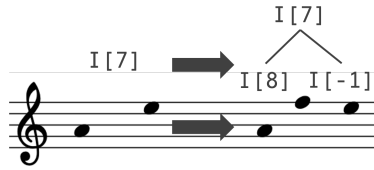


図 4.6: 非終端記号導出の様子

4.3.2 パラメータ推定

本研究では, 次の 3 種類の分布をコーパスから学習する.

1. ある非終端記号が 1 つの導出規則を選択する際にしたがう確率 θ
2. ある非終端記号が終端記号を導出する確率 λ
3. 既知の θ, λ にしたがって, 終端記号列 w を導出できる木 t (本研究のように, 対象に文法曖昧性が存在する場合, 1 つの w を導出する木は複数存在する)

その方法として, ギブスサンプリングによる事後分布推定の枠組み [45] を用いた. 以下にその詳細を述べる.

PCFG のパラメータ推定に用いられるギブスサンプリングでは, 終端記号列の集合 \mathbf{w} が与えられた時の事後分布 $P(\theta|\mathbf{w})$ を推定できる. そのために, 導出規則選択確率 θ と終端記号出力確率 λ , 終端記号列 w の導出木 t の集合 \mathbf{t} を次の式に従って交互にサンプリングする.

$$P(\mathbf{t}|\mathbf{w}, \theta, \lambda) = \prod_{i=1}^n P(t_i|w_i, \theta, \lambda) \quad (4.5)$$

$$\begin{aligned} P(\theta|\mathbf{t}, \mathbf{w}, \alpha) &= P_D(\theta|\mathbf{f}(\mathbf{t}) + \alpha) \\ &= \prod_{A \in N} P_D(\theta_A|f_A(\mathbf{t}) + \alpha) \end{aligned} \quad (4.6)$$

$$\begin{aligned} P(\lambda|\mathbf{t}, \mathbf{w}, \iota) &= \text{Beta}(\lambda|\mathbf{g}(\mathbf{t}) + \iota) \\ &= \prod_{A \in N} \text{Beta}(\lambda_A|g_A(\mathbf{t}) + \iota_A) \end{aligned} \quad (4.7)$$

P_D はディリクレ分布を表す. α, ι は任意のベクトルであり, 学習時にハイパーパラメータとして与える. $f_A(\mathbf{t})$ とは, サンプリグされた \mathbf{t} 中に存在する A を左辺とする導出規則の数を, 右辺毎に数えあげて, ベクトルとして返す関数である.

$\mathbf{g}_A(\mathbf{t})$ とは, \mathbf{t} 中の非終端記号 A について, 終端記号を出力している回数と非終端記号を出力している回数を数えあげ, ベクトルとして返す関数である.

式 (4.5) の木のサンプリグは, 次に示す内側フィルタリングを用いることで効率的に行う. ある終端記号列 $w = (w_0, w_1, \dots, w_n)$ から 1 つの木 t をサンプリグする. そのために, 終端記号列の部分列 $(w_i, w_{i+1}, \dots, w_j)$ が非終端記号 A から生成される確率を内側確率 $p_{i,j}^A$ として定義する. この内側確率は, 次の式に従う.

$$p_{i,i}^A = \lambda_A \quad (4.8)$$

$$p_{i,j}^A = \sum_{A \rightarrow BC \in R} \sum_{i \leq k < j} (1 - \lambda_A) \theta_{A \rightarrow BC} p_{i,k}^B p_{k+1,j}^C \quad (4.9)$$

$\theta_{A \rightarrow BC}$ は A が非終端記号 BC を生成する確率を表す. $p_{i,j}^A$ を $A \in N, 0 \leq i < j \leq n$ の範囲にわたって求めることで, 木のサンプリグが可能となる.

サンプリグ関数の擬似コードを 4.1 に示す.

4.1: 木サンプリグの手順

```

1 function Sampling(A, i, j)
2     if i=j then return Tree(A)
3     else
4         (k, B, C) = Multi(A, i, j)
5         return Tree(A, Sampling(B, i, k), Sampling(C, k+1, j))

```

関数 `Multi` は以下の確率で非終端記号 B, C と終端記号列のインデックス k を返す関数とする.

$$P(k, B, C) = \frac{(1 - \lambda_A) \theta_{A \rightarrow BC} p_{i,k}^B p_{k+1,j}^C}{p_{i,j}^A} \quad (4.10)$$

第5章 結果と考察

本章では、学習によって得られた以下の結果について述べる。

- サンプルされた木で選択された導出規則の集計
- 導出規則の分布 (θ_N)
- 各非終端記号からの終端記号の出力確率 (λ_N)

サンプルされた木と θ_N ($N = \{I[-16], I[-15], \dots, I[16]\}$) をもとに、導出規則の出現回数、各 θ の分布を図示し、適宜考察を与える。

また、学習に用いたセグメント中で、本研究で定義した文法規則では導出不可能となる例が存在した。それらについても取り上げ、導出不可能となった理由をまとめる。

5.1 サンプリング結果

5.1.1 頻度分析

バップのセグメント、モードのセグメントそれぞれのサンプリングによって出現した導出規則の回数 (上位のみ) を図 5.1 と図 5.2 に示す。

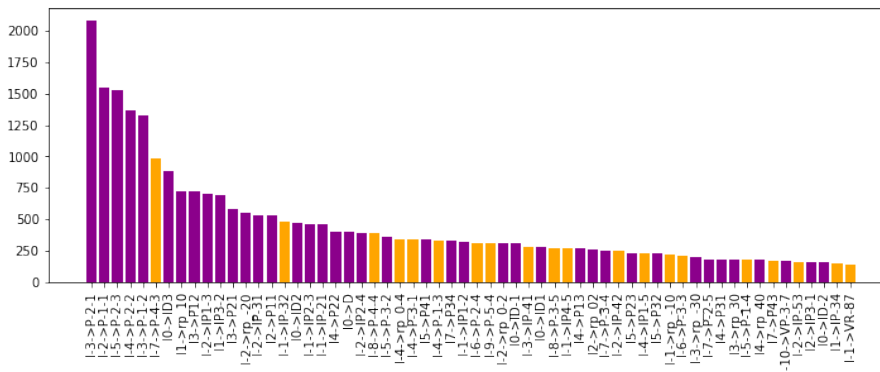


図 5.1: バップにおける導出規則上位 60

図 5.1 は、バップのセグメントにおいて各導出規則が木のサンプル中で何回使われたかの集計である。出現回数上位 60 位を表示している。縦軸が出現回数を、横軸が導出規則のラベルを表す。導出規則のラベルは 4.3.1 節で定義した文法 R の左辺と右辺を \rightarrow で区切り表現している。例えば、ラベル「1-3->P-2-1」は、「I[-3] において P 規則が選択され、I[-2] と I[-1] が導出された」ことを意味する。また、燈色の棒は、バップの上位 60 位には出現したが、モードの上位 60 位には含まれない導出規則であることを示す。

上位 6 位は下行方向の音程に関する導出が集中した。P や IP といった狭い音程のみで構成される導出規則が目立つ。

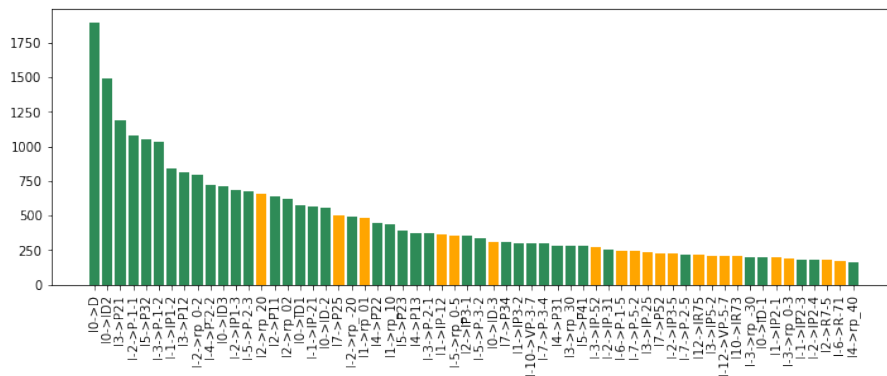


図 5.2: モードにおける導出規則上位 60

図 5.2 は、モードのセグメントを導出する木のサンプルに選択された導出規則の集計である。出現回数上位 60 位を表示している。燈色の箇所は、モードの上位 60 位には出現したが、バップの上位 60 位には入っていない導出規則であることを示す。上位 2 つは左辺が I0 の導出規則となった。また、バップの結果と比べて *repeat* 規則の出現が目立つ。これらより、同音連打が多く存在したと推測できる。

バップ上位規則における傾向

バップの導出規則では、下行する音形に関する規則が上行形より多く存在した。バップには狭い音程を続けることで下行し、広い音程で一気に上行する¹形の音形が多く存在することが影響したと考えられる。学習した即興演奏中の音程をヒストグラムで表示したものが図 5.3 と図 5.4 である。実際、バップのヒストグラムにおける $-5 < n < 5$ の範囲はモードのヒストグラムに比べ非対称であり、負の音程の出現頻度がモードより高いことがみてとれる。また、図 5.3 の、+7 から右の範囲は小さいがピークができています。これは広い音程で一気に上行する箇所の存在を意味する。

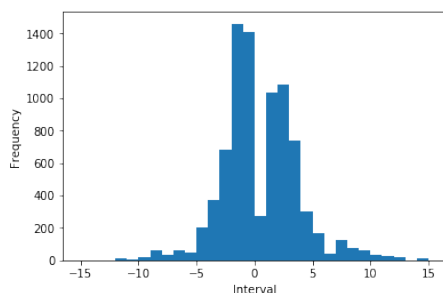


図 5.3: バップの音程集計

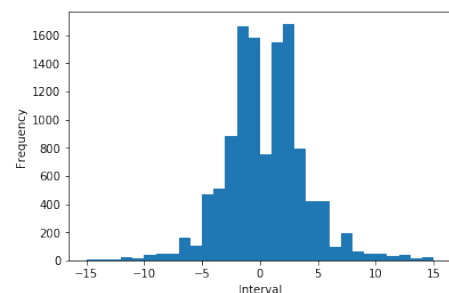


図 5.4: モードの音程集計

非終端記号の出現数においても、バップはモードに比べて下行形のものが多い。確認のために、サンプル中の非終端記号の使用回数を図 5.5 に示す。また、非終端記号別の終端記号出力確率 $\lambda_{I[n]}$ を図 5.6 に示す。横軸は非終端記号の種類、縦軸がその出現回数を示す。バップの値を紫色、モードの値を緑色で表している。

¹ 上行下行の前後関係は結果から判断できない

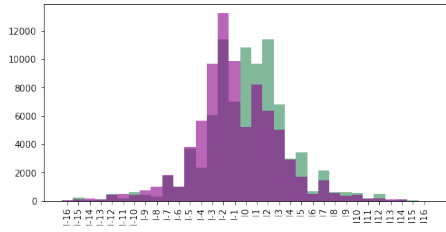


図 5.5: 非終端記号の出現回数

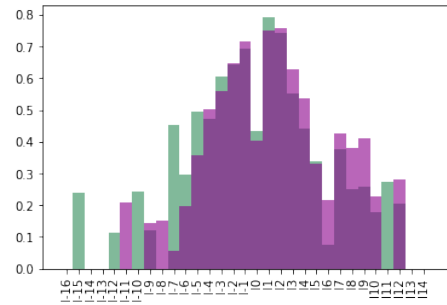


図 5.6: $\lambda_{I[n]}$

図 5.5 に紫色で表されるバップの非終端記号使用回数は、 $I[-2]$ が最も高く、 $I[-3]$ や $I[-1]$ も高くなっていることがわかる。緑色で示されたモードの分布形状と比べると、ピークが左に存在する。このことから、バップにおける下行音程は狭くなるという傾向が言える。

また、図 5.6 においては、バップにおける「広い音程を用いて上行する」傾向が見られる。バップ (紫色) の $I[7]$ より右側の出力確率を見ると、モードの確率や、反対側の広い音程 ($I[-7]$ から左側) より高くなっている。これより、バップにおいては、広い音程での上行が、演奏上で隣接する音符間で現れる可能性がモードより高いと言える。

モード上位規則における傾向

一方、モードの導出規則上位 (図 5.2) には、*repeat* や D , ID といった、音程が 0 に関する導出規則がバップに比べ多く現れた。これは、同音連打や同音高の音符で挟まれた音形が導出されやすいことを意味する。その理由として、モードにおける、停滞感や特性音の強調を目的とした同音の連続使用が学習されたからと考える。

さらに、グラフの傾斜に注目するとバップのものとは比べなからである。50 位あたりには、バップには見られない R や IR に関連する規則も出現している。これらよりバップの導出規則は狭い音程をさらに分割する導出規則が集中して出現していたことに対し、モードの導出規則の出現率はバップに比べ分散しているといえる。比較のために、導出規則の使用回数上位 60 位を、関連する IRM シンボルごとに集計し、割合にまとめたものを表 5.1 に示す。

表 5.1: バップとモードの上位規則におけるシンボル別使用回数割合

	P	IP	VP	R	IR	VR	D	ID	<i>repeat</i>
バップ	58.4	20.4	0.6	0	0	0.5	1.4	7.9	10.9
モード	41.5	17.2	1.8	1.2	1.5	0	6.7	13.5	16.4

割合を比較すると、バップにおける導出規則は P と IP (どちらの規則も狭い音程 2 つを導出する) が 8 割を占める。一方モードは広い音程を含む音形を導出する規則のうち VP , R , IR が (残りは VP) バップより高くなっていることがわかる。また、同音高を含む音形を導出する規則 $D, ID, repeat$ においても割合が高い。

バップとモードにおける固有の音形

図 5.1 と図 5.2 に橙色で示した導出規則に注目する。これらの規則は互いの上位には出現していない導出規則である。つまり、橙色の規則はバップ／モードそれぞれに固有の音形を導出する規則であると考えられる。

ここでバップに存在する $I[-7] \rightarrow I[-4] I[-3]$ やモードに存在する $I[-7] \rightarrow I[-5] I[-2]$ といった、5半音または7半音に関連する導出規則に着目する。5半音または7半音は、完全4度または完全5度に相当する。この2種類の音程は、周波数比がほぼ3:2となる音程を形成する。また、完全5度は音楽構造の認知に重要な働きをもつ [46]。

バップに存在する $I[-7] \rightarrow I[-4] I[-3]$ は、短三和音を5度3度1度と下行する動き、または長七の和音を7度5度3度と下行する動き (図 5.7 参照) と一致する。つまり、人間が和声に基づいて演奏する時と同じような導出規則を学習したといえる。

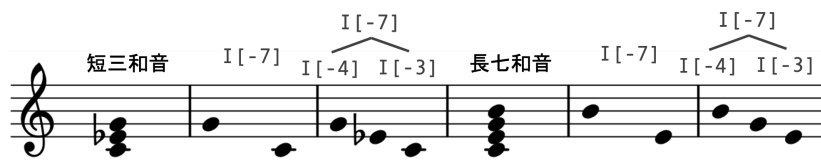


図 5.7: バップの特徴的導出規則の考察

一方、モードの橙色ラベルには $I[7] \rightarrow I[2] I[5]$ 、 $I[-7] \rightarrow I[-5] I[-2]$ といった導出規則が存在する。このような5半音を含む動きでは、三和音のような基本的な和音を連想しづらい。また、モードにおける上位の導出のうち、バップの上位に存在しない規則をまとめると (表 5.2)、絶対値が5となる音程が右辺によく見られることがわかる。5半音の音程は完全4度であり、核音²を構築する。つまり、モードはバップに比べて、核音に基づくような観点から構文解析されたと言える。

表 5.2: モード導出規則の上位とバップ導出規則上位の差集合

左辺	右辺
$I[2]$	$I[2] I[0]$ $I[7] I[-5]$
$I[7]$	$I[2] I[5]$ $I[5] I[2]$
$I[1]$	$I[0] I[1]$ $I[-1] I[2]$ $I[2] I[-1]$
$I[-5]$	$I[0] I[-5]$
$I[0]$	$I[-3] I[3]$
$I[-3]$	$I[-5] I[2]$ $I[0] I[-3]$
$I[-6]$	$I[-1] I[-5]$ $I[-7] I[1]$
$I[-7]$	$I[-5] I[-2]$
$I[3]$	$I[-2] I[5]$ $I[5] I[-2]$
$I[-2]$	$I[3] I[-5]$
$I[12]$	$I[7] I[5]$
$I[-12]$	$I[-5] I[-7]$
$I[10]$	$I[7] I[3]$

²民謡において旋律を構成する大枠となる [47]。また、和声整備以前の西洋音楽においても重要である

5.1.2 IRM シンボル別分析

導出規則の違いを詳細に比較するために、導出規則に対応する IRM シンボル別にサンプルを集計した結果を図 5.8, 図 5.9, 図 5.10, 図 5.11 に示す。縦軸が出現回数を、横軸が導出規則のラベルを表す。燈色は、一方の上位に入ったが他方の上位には入らなかった導出規則を示す。

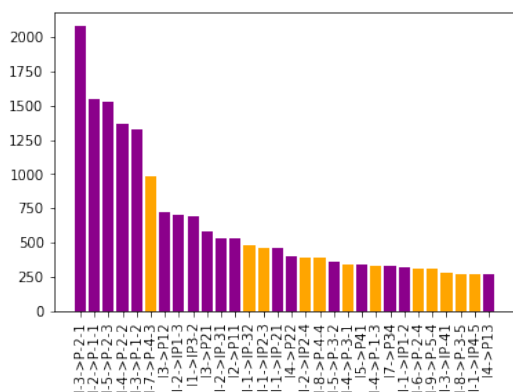


図 5.8: バップの P 系導出規則上位 30

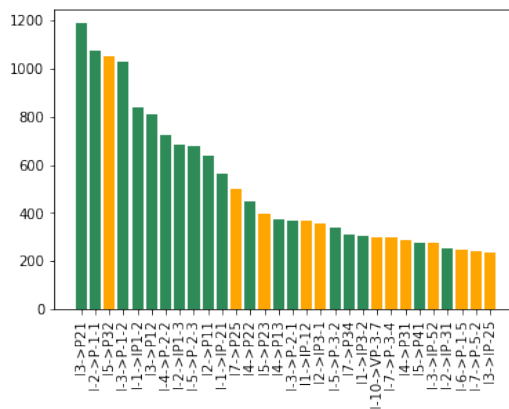


図 5.9: モードの P 系導出規則上位 30

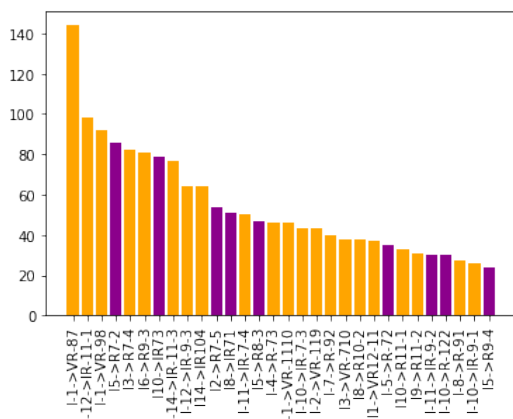


図 5.10: バップの R 系導出規則上位 30

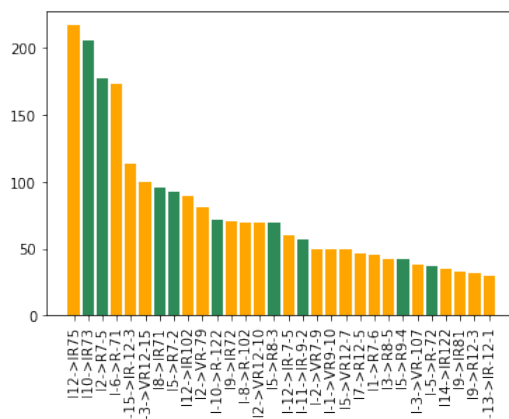


図 5.11: モードの R 系導出規則上位 30

考察

P 系の導出規則を集計した結果、図 5.9 の最後の導出規則 $I[3] \rightarrow I[-2]I[5]$ を除き全体の上位 60 位にランクインしていることがわかった。つまり、セグメントの大部分は、妥当な進行を感じさせるような音形となりやすい。これは、意外な展開は稀に現れることに意味があるような音楽の性質と合致している。また、本研究で分析したセグメントは短いひとまとまりの旋律である。つまり、セグメント内で終止や意外な展開を感じることは少ないといえる。この観点からも、セグメント内には終止を感じさせることのできる R 系ではなく、進行を続けると予感させる P 系で構成されることには妥当性がある。

バップ特有の R 系導出規則

R 系の導出規則の出現傾向は、バップとモードの間で大きく異なった。これより、R 系の導出規則は、出現回数は多くないものの、バップまたはモード固有の音形に対する学習をしたと考える。実際にどのような音形と解析木が導出されるかの例を以下に示す。

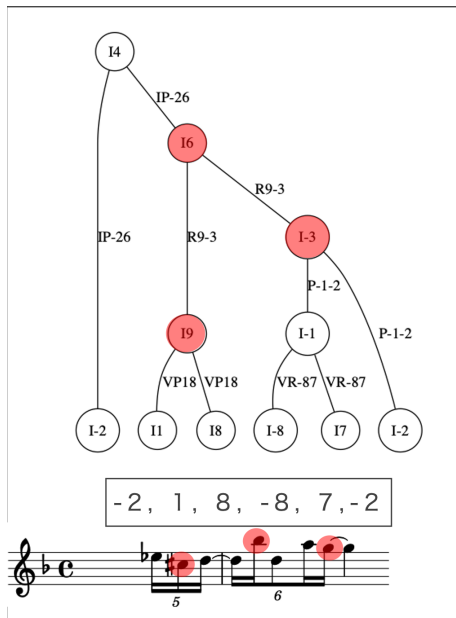


図 5.12: Charlie Parker, Embraceable You m.19

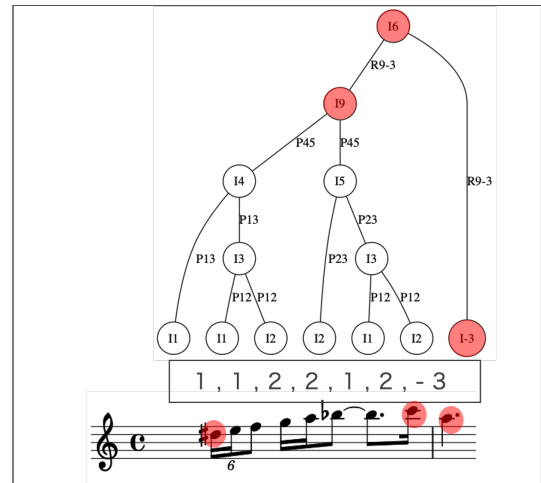


図 5.13: Charlie Parker, Don't Blame Me m.1

図 5.12 と図 5.13 は、木中にバップ特有の導出規則 $I[6] \rightarrow I[9] I[-3]$ が現れたセグメントの例である。木構造と音程系列，譜例を同時に示す。この導出規則に関する共通点が 2 つ存在する。

1. 導出される非終端記号 $I[9]$ は、セグメント中の最低音と最高音の音程である。
2. 導出される非終端記号 $I[-3]$ は、セグメント中の最高音とセグメントの最終音の音程である。

これより、導出規則 $I[6] \rightarrow I[9] I[-3]$ は、セグメント概形に影響する導出規則であると考えられる。そして、この導出規則がバップの上位のみ出現することより、バップ特有の概形である可能性も示唆される。実際、 $I[9]$ と $I[-3]$ の音程関係は 5.1.1 節にて述べた「上行は広い音程，下行は狭い音程」というバップの全体的な傾向を端的に表している。また、減七和音を構成するような音程 (3, 6, 9 半音) の導出規則となっていることも興味深い。減七和音やそれに基づく音階は、機能 and 声における緊張の働きが強く、バップの即興演奏で多用される。



図 5.14: 減七和音とそれに基づく音階

モード特有の R 系導出規則

モードに出現した R 導出規則に基づくセグメントについても例を示す。

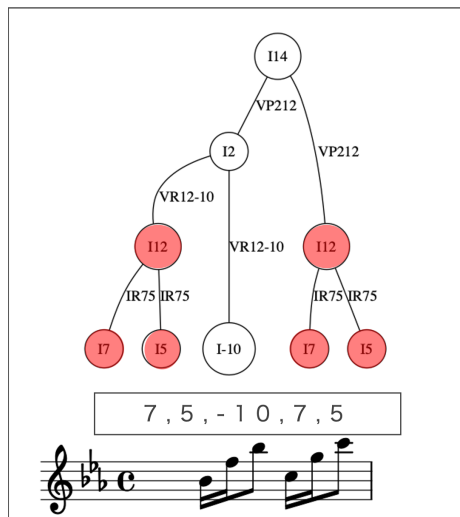


図 5.15: Bob Berg, Nature Beast, m.63

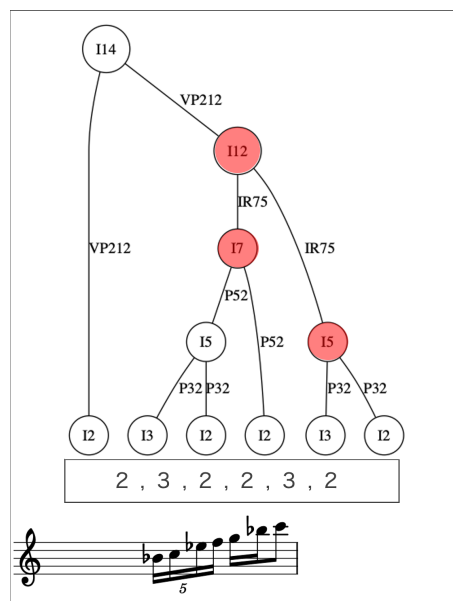


図 5.16: Kenny Wheeler, Pass It On, m.16

図 5.15 と図 5.16 は、木中にモード特有の導出規則 $I[12] \rightarrow I[7] I[5]$ が現れたセグメントの例である。これらのセグメントは、それぞれ Bb メジャーペンタトニックスケールと C マイナーペンタトニックスケールで構成されている。ペンタトニックスケールとは、鍵盤の黒鍵の音程関係からなる音階である。ペンタトニックスケールの 5 音は、半音で隣接することがなく、どの音の組み合わせをとっても調子はずれにならない。これらの例では、少ない音数で安定した調性を感じさせるためにペンタトニックスケールを用いたと推測できる。つまり、モード特有の導出規則には、旋律の構成音が協和することを重要視するものがあると確認できた。別の例として、図 5.15 と図 5.16 に出現する $I[14] \rightarrow I[2] I[12]$ や $I[2] \rightarrow I[12] I[-10]$ においても、もっとも協和する音程であるオクターブを構成できる導出規則を選択していることがうかがえる。

5.1.3 各非終端記号における導出規則選択確率

非終端記号別の導出規則選択確率を本節にて示す。ここでは考察をおこなった非終端記号についてのみ触れ、全ての結果については付録に示す。

I[±7] の導出規則選択確率

完全5度の音程に対応する非終端記号 I[7] と I[-7] についての考察を述べる。

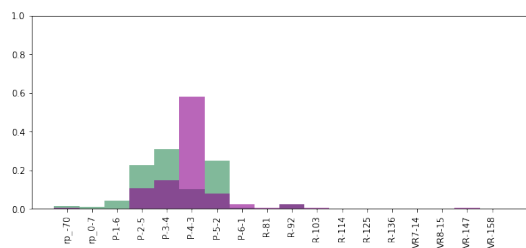


図 5.17: $\theta_{I[-7]}$

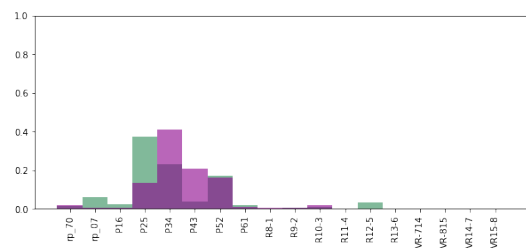


図 5.18: $\theta_{I[7]}$

5.1.1 節で考察したように、バップ(紫色)は2つの3度音程へ分割する規則 (P-3-4, P-4-3, P34, P43) の選択確率が高く、モード(緑色)は完全4度と1度に分割する規則 (P-2-5, P-5-2, P25, P52) の選択確率が高い。

バップ間の $\theta_{I[-7]}$ と $\theta_{I[7]}$ において、導出後の非終端記号の音程の絶対値が4と3になる規則を比較する、 $\theta_{I[-7]}$ では $\theta_{I[-7] \rightarrow I[-4]I[-3]} > \theta_{I[-7] \rightarrow I[-3]I[-4]}$ と、4が前になる確率は、4が後になる確率より高い。一方 $\theta_{I[7] \rightarrow I[4]I[3]} < \theta_{I[7] \rightarrow I[3]I[4]}$ であり、順序が逆転することがわかる。これは、短三和音は短3度 → 長3度が堆積して構成されており、下行する場合は4半音 → 3半音となるが、上行する場合は3半音 → 4半音となるような非対称性と辻褄が合う。つまり、 $\theta_{I[-7]}$ と $\theta_{I[7]}$ は、人間が音楽的構造(短三和音や音階など)に則って演奏する時と類似している。モードの導出規則においても、同様に順序の逆転関係が言える。

I[±12] の導出規則選択確率

オクターブの音程 $\theta_{I[-12]}$ と $\theta_{I[12]}$ では、符号間やバップとモードの間で分布に大きな違いが現れた。

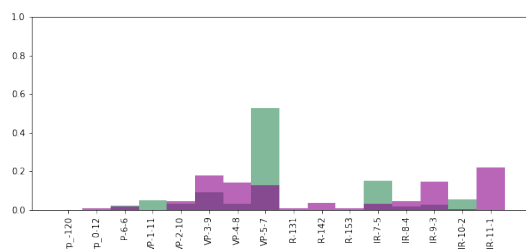


図 5.19: $\theta_{I[-12]}$

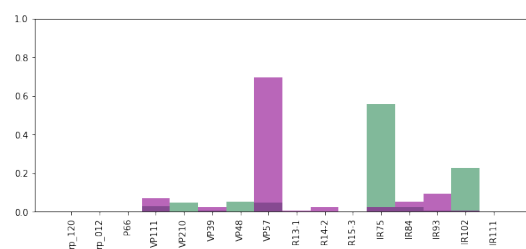


図 5.20: $\theta_{I[12]}$

バップ同士の $\theta_{I[-12]}$ と $\theta_{I[12]}$ 比較 バップにおける $\theta_{I[-12]}$ は分布がばらけている．それに対し $\theta_{I[12]}$ では $I[5]I[7]$ を導出する確率に集中している． $\theta_{I[-12]}$ において分布が平坦な理由は、オクターブはいかなる和音にも分割可能な音程関係であることと考える．つまり、12半音を和音の構成音に従って分割しようとする際、長七和音(1と11)、短七和音(2と10)、減七和音(3と9)、増三和音(4と8)、完全五度の転回(5と7)のように、全ての組み合わせで基本的な和音に関連づけることができってしまう．このようなケースが学習で生じた結果、 $\theta_{I[-12]}$ において分布が平坦になったと考える．

一方、上行形では大きい音程は分割されず、終端記号として出力されやすい。「オクターブ上行の音程は $I[5]I[7]$ に分割されたのち、 $I[7]$ が終端記号を出力する」という導出規則の選択確率が高いことで広い音程で上行するバップの傾向が観測されると推測する．

モード同士の $\theta_{I[-12]}$ と $\theta_{I[12]}$ 比較 モードにおける導出規則では、7半音と5半音に分割する導出規則で確信度が高くなった．下行と上行で確信度が最大となる VP-5-7 と IR75 において $I[\pm 7]$ で考察した順序の逆転が見られる．この音程の組み合わせは、例えば C-G-C のような音高の組み合わせ(音楽における代数関係で置き換えると I-V-I と表現される)になる．この2種類の音では、c が安定、g が緊張と感じるとされる．もし $I[-12]$ を $I[-7]I[-5]$ のように分割した場合は、音高の組み合わせが G-C-G(V-I-V) となり、緊張を多く含むことになる．つまり、モードでは協和する響きの非終端記号を導出し、順序においても安定性が高くなる組み合わせを選択しやすくなっている．

$\theta_{I[-12]}$ のバップとモード比較 $\theta_{I[-12]}$ におけるバップとモードの導出傾向を比較する．バップはあらゆる音程の組み合わせの可能性がある．一方モードは協和する音程を中心に選択する．

$\theta_{I[12]}$ のバップとモード比較 $\theta_{I[12]}$ では、バップもモードも協和する音程の $I[5]$ と $I[7]$ を導出しやすい．しかし順序を考慮するとバップが V-I-V を導出し、モードは I-V-I を導出するという違いが存在する．

5.2 文法から導出不可能なセグメント

本節では導出不可能(全ての非終端記号において、開始記号である確率が0)となったセグメントについて述べる．これらのセグメントは学習から除外された．

5.2.1 バップの導出不可能セグメント

まず、バップの導出不可能なセグメントについて述べる．



図 5.21: Charlie Parker, Don't Blame Me, m. 11. 図 5.22: Charlie Parker, Embraceable You, mm. 22-23

バップのセグメントにおいて、本研究の文法で導出不可能なものは2つ存在した．図 5.21 に示す Charlie Parker 演奏の Don't Blame Me の 11 小節目に存在するセグメントと、図 5.22

に示す Charlie Parker 演奏の Embraceable You の 22 小節目と 23 小節目にまたがるセグメントである。それぞれについて導出不可能となった理由をのべる。

図 5.21 このセグメントは最初の音 D と最後の音 A の音程が 17 半音になる。本研究で非終端記号として扱える音程の上限 16 半音をこえたため導出不可能となった。

このセグメントはひとつひとつの音程の推移に例外性は低く、上限を広げればその他のセグメントと同様導出可能となる。

図 5.22 このセグメントは最初の Eb と Bb が 7 半音の音程にある。この音程は本研究では「広い音程」とされる。IRM の基本原理に基づくと、次の音程は、「狭い音程」または「逆方向の音程」どちらかの暗意が実現されることを期待する。しかし、次の Bb と G の作る音程は「広い音程」で「同方向」である。さらに後ろの F まで考慮してもその条件は変わらない。つまり、このセグメントは IRM の基本原理による暗意と実現の構造を満たさないので、本研究の文法からは導出不可能であった。

5.2.2 モードの導出不可能セグメント

続いて、モードに存在した導出不可能なセグメントを図 5.23–図 5.28 に示す。モードはバップより多く、9 個 8 種類のセグメントが導出不可能であった。





図 5.26: George Coleman, Maiden Voyage, mm.13-15.

図 5.26 は George Coleman 演奏の Maiden Voyage に出現したセグメントである。この譜例は、休符で隔たれた 3 つのセグメントが連続している。図 5.26 に含まれる 3 つのセグメント全てが導出不可能となった。その原因は共通して存在する $Bb \rightarrow F \rightarrow C$ の音形である。この音形も、本研究では「狭い音程 \rightarrow 逆方向の広い音程」という未定義の音形に分類されるため導出不可能となる。



図 5.27: Chris Potter, PopTune, m. 50.



図 5.28: Chris Potter, PopTune, mm.110-111.

図 5.27, 図 5.28 は、Chris Potter 演奏の PopTune に出現したセグメントである。これらのセグメントも、「狭い音程 \rightarrow 逆方向の広い音程」で定義される文法が導出に必要である。導出不可能の理由は、以下の 3 つにまとめることができた。

1. 非終端記号の音程上限を超える。
2. 「広い音程」 \rightarrow 「同方向の広い音程」を許す導出規則が必要。バップに 1 箇所
3. 「狭い音程」 \rightarrow 「逆方向の広い音程」を許す導出規則が必要。モードに 7 箇所

2. に分類されたセグメントは上行する形であった。つまりバップの傾向として考察される「広い音程での上行」を極端にした例といえる。また、3. に分類されるセグメントも 5 半音と 7 半音で構成される動きであり、「協和する音程となる非終端記号の導出」というモードの傾向が極端に表出している例である。

第6章 終章

6.1 まとめ

本研究には、i) 即興演奏中の構造を解析するモデルの構築、ii) バップとモード間の構造の差異を抽出する、という2段階の目的が存在した。

計算機における音楽の理解を目的とした研究は盛んに行われている。その中でも、記号表現された音楽を解析することは重要といわれている。本研究では、そのような立場に基づき、「確率的音楽生成モデルを構築し、教師なし学習によって階層的構造と、その構造を導く規則を学習する」手法を選択した。本研究独自の観点として、i)IRMの原理に基づく確率文脈自由文法モデルを構築したこと、ii)機能と声に基づくバップと、旋法に基づくモードを同じ観点から分析することの2点が挙げられる。人間の音楽認知におけるIRMを取り入れることで、学習結果から得られた構造がどのような情動と結びつくかを考察できる。また、バップとモードの対比は、音楽における普遍要素を考察することを狙いとした。

そして、IRMにおける暗意の原理をもとに、音程の系列を導出する文法を定式化した。さらに、定義した文法に基づく確率的文脈自由文法モデルを構築した。その後、バップとモードそれぞれを対象として構築したモデルの学習をおこなった。

学習結果から、i)バップとモード間で共通して多数使用される導出規則の傾向、ii)バップのみ、モードのみで多数使用される導出規則の傾向、iii)IRMにおけるP系とR系に関連する導出規則の傾向、iv)非終端記号別の導出規則の傾向に着目した。それぞれにおいて、次のような考察を得た。

- i) P系の音形を導出する規則が使用回数の大部分を占める。共通する導出規則においても、バップは「狭い音程で下行し、広い音程で上行する」ような旋律を導出する規則が多く使用され、モードは同音連打を含む旋律を導出する規則が多く使用されている。
- ii) バップでのみ多数使用された導出規則には、完全5度を3度に分割する(または3度音程を積み上げて完全5度を形成する)、人間が和音に基づいて演奏する際と同じような規則がみられた。一方、モードでのみ多数使用された導出規則には、完全4度を旋律の枠組みとするような規則がみられた。
- iii) P系の導出規則ではi)と同様の傾向が確認できた。R系の規則では、バップとモード間で使用傾向に大きな差が現れた。特にモードでは、完全4度や完全5度の協和する音程に関連する規則が学習された。
- iv) 上行音程に対応する非終端記号と、下行音程に対応する非終端記号の間で導出規則に対称性が確認できた。この対称性も、人間が音楽理論をもとに演奏すると現れる。また、完全5度やオクターブのような協和音程を2種類の音程に分割する際に、バップにおいては和音に含まれる音程を導出する傾向が強く、モードにおいては協和する音程を優先して導出する傾向が強くなった。

6.2 今後

本研究の応用として、即興演奏の構造分析システムの構築が考えられる。例えば、与えられた即興演奏の構文解析を行い、結果を楽譜上に図示することで、階層構造や、旋律の骨格をなす音程を分析できるシステムである。そのようなシステムを構築することによって、即興演奏の分析や理解を助けることを期待する。

本研究が残した課題として、完全な楽譜を生成するには不十分である点が挙げられる。本研究を発展させる今後の方針として、まずは音程系列を入力として、音の開始時刻や音長を出力するようなリズム生成モデルを構築する。リズム生成モデルの構築後に本研究で構築したモデルとの階層化を行い、完全な楽譜を生成可能にする。

階層化したモデルが完成すれば、さらなる応用が可能となる。例えば、編曲や作曲、即興演奏支援システムへの応用が考えられる。また、モデルが生成した即興演奏を人間が評価することによって、モデルの妥当性やIRMによる旋律簡約の理論構築の方針を検討できるようになる。

学習法に関しても発展の余地がある。まず、学習速度に関しては、木構造をサンプルする際の内側確率の計算が最も時間のかかるプロセスであった。この計算時間は「音程系列の長さ³ × 非終端記号の数³」に比例する。この問題に対し、計算時間を削減する手法 [48] が提案されている。

本研究で設定した導出規則では、導出不可能となるセグメントが存在した。その問題に対する改善も今後の課題としたい。例えば、Namour の提唱した Retrospective 形のシンボルを取り入れることで、導出不可能の問題は解消される。また、今回の文法における導出規則と非終端記号は手作業で設定したが、木構造生成モデルのノンパラメトリック化 [49][50] によって非終端記号と導出規則を獲得させることも可能である。

謝辞

筑波大学の平賀譲先生には，本論文を執筆するにあたり，指導教員として多くの助言をいただきました。日頃の研究活動においても，様々な支援や示唆をいただきました。心から感謝いたします。私の筑波大学情報メディア創成学類在籍時から現在にわたり，筑波大学の寺澤洋子先生，松原正樹先生には，本研究分野に対する専門的な知見から，多くの情報提供や助言をいただきました。ありがとうございました。

また，同研究室の学生として，共に議論を深めあった，人と音の情報学研究室所属のみなさまに感謝いたします。

参考文献

- [1] 相倉久人. ジャズの歴史: 新書で入門. 新潮社, 2007.
- [2] カワイ音楽教育研究所. すぐに役立つ音楽用語ハンドブック: 音楽・教育・保育に携わる人々に. カワイ出版, 2013.
- [3] 田柳恵美子. 音楽のパフォーマンスデザインとイノベーション: ジャズにおける即興と革新を事例として. 認知科学, Vol. 17, No. 3, pp. 459–473, 2010.
- [4] 菊地成孔, 大谷能生, 飯野友幸. 東京大学のアルバート・アイラー: 東大ジャズ講義録キーワード編. 文芸春秋, 2009.
- [5] 北川祐. コード・スケールハンドブック ポピュラー音楽のためのスケール理論. リットーミュージック, 1995.
- [6] G.A. Russell. 調性組織におけるリディアン・クロマティック・コンセプト. エー・ティー・エヌ, 1993.
- [7] Eugene Narmour. *The analysis and cognition of basic melodic structures: The Implication-Realization Model*. University of Chicago Press, 1990.
- [8] 矢澤櫻子. 音楽理論暗意実現モデルに基づく楽曲解析に関する研究. PhD thesis, 筑波大学, 2016.
- [9] 吉井和佳. 2章 音楽と統計的記号処理. 映像情報メディア学会誌, Vol. 71, No. 4, pp. 457–461, 2017.
- [10] 吉井和佳ほか. 音楽を軸に広がる情報科学: 5. 音楽と機械学習. 情報処理, Vol. 57, No. 6, pp. 519–522, 2016.
- [11] David Cope. Computer modeling of musical intelligence in EMI. *Computer Music Journal*, Vol. 16, No. 2, pp. 69–83, 1992.
- [12] Damon Horowitz. Representing musical knowledge in a jazz improvisation system. *Proc. of Artificial Intelligence and Music*, pp. 16–23, 1995.
- [13] 和気早苗, 加藤博一, 才脇直樹, 井口征士ほか. テンション・パラメータを用いた協調型自動演奏システム: Jasper. 情報処理学会論文誌, Vol. 35, No. 7, pp. 1469–1481, 1994.
- [14] 後藤真孝, 平田圭二ほか. ハービー君: 演繹オブジェクト指向に基づいてジャズらしいコードにリハーモナイズするシステム. 情報処理学会研究報告音楽情報科学 (MUS), Vol. 1996, No. 75 (1996-MUS-016), pp. 33–38, 1996.
- [15] Sakurako Yazawa, Masatoshi Hamanaka, and Takehito Utsuro. Melody generation system based on a theory of melody sequence. In *2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA)*, pp. 336–341. IEEE, 2014.

- [16] Jon Gillick, Kevin Tang, and Robert M Keller. Machine learning of jazz grammars. *Computer Music Journal*, Vol. 34, No. 3, pp. 56–66, 2010.
- [17] S Sertan and Parag Chordia. Modeling melodic improvisation in Turkish folk music using variable-length Markov models. In *12th International Society for Music Information Retrieval Conference*, pp. 269–274. Citeseer, 2011.
- [18] Martin Rohrmeier. A generative grammar approach to diatonic harmonic structure. In *Proceedings of the 4th sound and music computing conference*, pp. 97–100, 2007.
- [19] Daniel Harasim, Timothy JO ’ Donnell, Martin Rohrmeier. Harmonic syntax in time rhythm improves grammatical models of harmony. 2019.
- [20] Robert M Keller and David R Morrison. A grammatical approach to automatic improvisation. Proceedings, Fourth Sound and Music Conference, Lefkada, Greece, July. “Most of the soloists at Birdland had to wait for Parker ’ s next record in order to find out what to play next. What will they do now, 2007.
- [21] Fred Lerdahl and Ray S Jackendoff. *A Generative Theory of Tonal Music*. MIT press, 1996.
- [22] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo. Melody morphing method based on GTTM. In *ICMC*, pp. 155–158. Citeseer, 2008.
- [23] Eita Nakamura, Masatoshi Hamanaka, Keiji Hirata, and Kazuyoshi Yoshii. Tree-structured probabilistic model of monophonic written music based on the generative theory of tonal music. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 276–280. IEEE, 2016.
- [24] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo. σ GTTM III: Learning-Based Time-Span Tree Generator Based on PCFG. In *International Symposium on Computer Music Multidisciplinary Research*, pp. 387–404. Springer, 2015.
- [25] John A Biles, et al. GenJam: A genetic algorithm for generating jazz solos. In *ICMC*, Vol. 94, pp. 131–137, 1994.
- [26] George Papadopoulos and Geraint Wiggins. A genetic algorithm for the generation of jazz melodies. *Proceedings of STEP*, Vol. 98, , 1998.
- [27] Kyoko Komatsu, Tomomi Yamanaka, Masami Takata, and Kazuki Joe. A music composition model with genetic programming. In *PDPTA*, pp. 686–692, 2010.
- [28] Mark Mcfarland. Schenker and the tonal jazz repertory: A response to Martin. *Music Theory Online*, Vol. 18, , 09 2012.
- [29] Steve Larson. Schenkerian analysis of modern jazz: questions about method. *Music Theory Spectrum*, Vol. 20, No. 2, pp. 209–241, 1998.
- [30] Henry Martin. *Charlie Parker and thematic improvisation*. Scarecrow Press, 1996.
- [31] Gunther Schuller. Sonny Rollins and the challenge of thematic improvisation. *The Jazz Review*, Vol. 1, No. 1, pp. 6–21, 1958.

- [32] Rafael Ramirez, Amaury Hazan, Esteban Maestre, and Xavier Serra. A genetic rule-based model of expressive performance for jazz saxophone. *Computer Music Journal*, Vol. 32, No. 1, pp. 38–50, 2008.
- [33] Belinda Thom. Unsupervised learning and interactive jazz/blues improvisation. In *AAAI/IAAI*, pp. 652–657, 2000.
- [34] Jon Snyder and Marti Hearst. ImproViz: visual explorations of jazz improvisations. In *CHI'05 extended abstracts on Human factors in computing systems*, pp. 1805–1808. ACM, 2005.
- [35] Martin Norgaard. How jazz musicians improvise: The central role of auditory and motor patterns. *Music Perception: An Interdisciplinary Journal*, Vol. 31, No. 3, pp. 271–287, 2014.
- [36] Klaus Frieler, Frank Höger, Martin Pfeiderer, and Simon Dixon. Two web applications for exploring melodic patterns in jazz solos. In *ISMIR*, pp. 777–783, 2018.
- [37] Dimitrios Vassilakis, Anastasia Georgaki, and Christina Anagnostopoulou. “jazz mapping” an analytical and computational approach to jazz improvisation. 2019.
- [38] Martin Pfeiderer and Klaus Frieler. The Jazzomat project. Issues and methods for the automatic analysis of jazz improvisations. *Concepts, experiments, and fieldwork: Studies in systematic musicology and ethnomusicology*, pp. 279–295, 2010.
- [39] Martin Pfeiderer, Klaus Frieler, Jakob Abeßer, Wolf-Georg Zaddach, and Benjamin Burkhart. Inside the Jazzomat. *New perspectives for jazz research. Schott Campus, Mainz, Germany*, 2017.
- [40] Christian Dittmar, Martin Pfeiderer, and Meinard Müller. Automated estimation of ride cymbal swing ratios in jazz recordings. In *ISMIR*, pp. 271–277, 2015.
- [41] Tetsuro Kitahara, Sergio Giraldo, and Rafael Ramirez. Jamsketch: Improvisation support system with GA-based melody creation from user’s drawing. In *International Symposium on Computer Music Multidisciplinary Research*, pp. 509–521. Springer, 2017.
- [42] Emilios Cambouropoulos. The local boundary detection model (LBDM) and its application in the study of expressive timing. In *ICMC*, p. 8, 2001.
- [43] Mark Johnson. PCFG models of linguistic tree representations. *Computational Linguistics*, Vol. 24, No. 4, pp. 613–632, 1998.
- [44] Édouard Gilbert and Darrell Conklin. A probabilistic context-free grammar for melodic reduction. In *Proceedings of the International Workshop on Artificial Intelligence and Music, 20th International Joint Conference on Artificial Intelligence*, pp. 83–94, 2007.
- [45] Mark Johnson, Thomas Griffiths, and Sharon Goldwater. Bayesian inference for PCFGs via markov chain monte carlo. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pp. 139–146, 2007.
- [46] Fred Lerdahl. *Tonal pitch space*. Oxford University Press, 2004.

- [47] 柴田南雄. 音楽の骸骨のはなし—日本民謡と 12 音音楽の理論. 東京: 音楽之友社, 1978.
- [48] 武井俊祐, 牧野貴樹, 高木利久. ベイズ確率文脈自由文法のための高速構文木サンプリング法. 情報論的学習理論テクニカルレポート, 2009.
- [49] Percy Liang, Slav Petrov, Michael Jordan, and Dan Klein. The infinite PCFG using hierarchical Dirichlet processes. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pp. 688–697, 2007.
- [50] Jenny Rose Finkel, Trond Grenager, and Christopher D Manning. The infinite tree. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 272–279, 2007.

付 録 A 分析に使用した曲

本研究の分析対象に選んだ曲を表 A.1 と表 A.2 に示す. 表 A.1 がバップの分析のため使用した曲, 表 A.2 がモードの分析のため使用した曲である.

表 A.1: 分析に用いた曲 (バップ)

曲名	演奏者	楽器	調	ノート数	BPM	録音年
Don't Blame Me	Charlie Parker	as	C-maj	516	63.7	1947
Donna Lee	Charlie Parker	as	A \flat -maj	371	225.7	1947
Embraceable You	Charlie Parker	as	F-maj	450	72.3	1947
How Deep Is the Ocean	Charlie Parker	as	C-min	375	71.9	1947
Ko-Ko	Charlie Parker	as	B \flat -maj	299.4	538	1945
My Little Suede Shoes	Charlie Parker	as	E \flat -maj	256	145.9	1951
Ornithology	Charlie Parker	as	G-maj	194	221	1946
Out of Nowhere	Charlie Parker	as	G-maj	482	67.6	1947
Scrapple from the Apple	Charlie Parker	as	F-maj	270	165	1947
Segment	Charlie Parker	as	B \flat -min	351	246.9	1949
Star Eyes	Charlie Parker	as	E \flat -maj	314	133.8	1950
Steeplechase	Charlie Parker	as	B \flat -maj	228	175.2	1949
Thriving on a Riff	Charlie Parker	as	B \flat -maj	357	227.2	1945
Yardbird Suite	Charlie Parker	as	C-maj	192	209.8	1946
Confirmation	Michael Brecker	ts	F-maj	1640	208.9	1981
I Mean You	Michael Brecker	ts	F-maj	551	180	1995
Airegin	Miles Davis	tp	A \flat -maj	431	241.3	1954
Oleo	Miles Davis	tp	B \flat -maj	224	271.1	1956
Oleo	Miles Davis	tp	B \flat -maj	223	268	1956
Tune Up	Miles Davis	tp	D-maj	409	281.5	1953

分析曲の表のうち, 曲名はアドリブの元となる曲の名前, 演奏者はアドリブを演奏している演奏者名を表す. 楽器は A.3 に示す略記を用いて表している. 調はデータベースに登録されている, 局所的な転調を無視した演奏全体の調を表示している. 登録のないものに関しては空白のままとした. ノート数とは, アドリブに出現する音符の総数を示す.

表 A.2: 分析に用いた曲 (モード)

曲名	演奏者	楽器	調	ノート数	BPM	録音年
Nature of the Beast	Bob Berg	ts	C-min	964	131.8	1993
So What	Cannonball Adderley	as	D-dor	445	142.2	1959
Arjuna	Chris Potter	ts	D-chrom	1318	117.7	2007
Pop Tune # 1	Chris Potter	ts	G-min	1353	108.6	2007
Rumples	Chris Potter	ts	F-min	821	134.9	2009
Togo	Chris Potter	ts	F-min	933	112	2007
Milestones	David Liebman	ss	F-maj	502	156	1992
Pablo's Story	David Liebman	ss	E-min	712	144.2	1988
Softly as in a Morning Sunrise	David Liebman	ss	C-min	1098	260.7	1988
Maiden Voyage	Freddie Hubbard	tp		552	126	1965
Maiden Voyage	George Coleman	ts		224	125.4	1965
Agitation	Herbie Hancock	p		343	293.8	1965
Pass it on	Kenny Wheeler	tp	C-min	264	105.7	1991
Agitation	Miles Davis	tp		432	252.8	1965
So What	Miles Davis	tp	D-maj	221	138.5	1959
Softly as in a Morning Sunrise	Milt Jackson	vib	C-min	844	137.3	1955
Pass it On	Steve Coleman	as		325	104.9	1991
Pass it On	Von Freeman	ts	C-min	461	105.9	1991

表 A.3: 楽器の略称記号と楽器名

ss	as	ts	tp	p	vib
ソプラノサククス	アルトサククス	テナーサククス	トランペット	ピアノ	ヴィブラフォン

付録B 分析時のパラメータの詳細

B.1 セグメンテーション時のパラメータ

分析データを分割する際に用いたLBDMでは、音高差, IOI, OOIそれぞれに上限 $Max_{pitch}, Max_{ioi}, Max_{ooi}$ を設定する。また、分割可能性の指標 s に3つの要素がどれだけ影響を与えるかの重み $w_{pitch}, w_{ioi}, w_{ooi}$ を設定する。本研究では、以下の値を用いた。

$$\begin{aligned} Max_{pitch} &= 12 \\ Max_{ioi} &= 3.0 \\ Max_{ooi} &= 3.0 \\ w_{pitch} &= 0.3 \\ w_{ioi} &= 0.4 \\ w_{ooi} &= 0.3 \end{aligned}$$

これらの値を使用し、 s が上位5%以上の値となる箇所を基準にセグメントへの分割を行った。

B.2 学習時のパラメータ

事前分布に与える α, ι

本研究では、確率的文脈自由文法の学習にベイズ事後分布推定の枠組みを使用した。多項分布となる導出規則選択確率の事前分布にディリクレ分布を、二項分布で表すことのできる終端記号出力確率の事前分布にベータ分布を設定した。事前分布と事後分布の共役性から事後分布を求めることが容易となる。

ディリクレ分布 $P_D(\theta|\alpha)$ は次の式で定義される。

$$P_D(\theta|\alpha) = \frac{1}{C(\alpha)} \prod_{i=1}^K \theta_i^{\alpha_i} \quad (\text{B.1})$$

$$C(\alpha) = \frac{\prod_{i=1}^K \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^K \alpha_i)} \quad (\text{B.2})$$

K は θ と α の次元を表す。また、 $0 \leq \theta_i \leq 1$, $\alpha_i > 0$ である。与えるパラメータ α によって、 θ の期待値が変化する。本研究では、IRMにて定義される暗意の強さ (図 B.1) に基づいて、ハイパーパラメータ $\alpha_{I[n]} (-16 \leq n \leq 16)$ を設定した。

暗意される音程が、先に聴こえた音程に対して (1) 似た性質¹(sim) なのか、(2) 遠い性質 (diff) なのか分類する。暗意の強さとは、その2つのどちらがより暗意されやすいかの関係を示したものである。

¹向きと大きさの2種類を合わせて性質と表現している

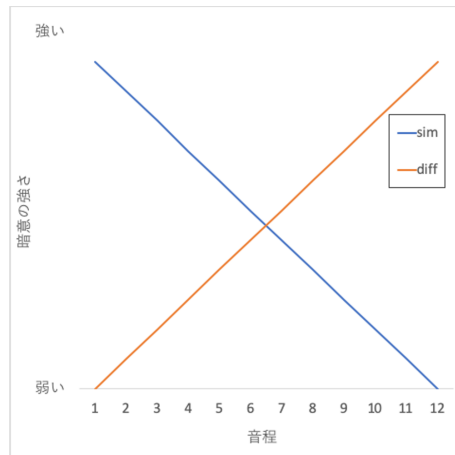


図 B.1: 暗意の強さの図式 (文献 [7]Fig.5.1 を参考に作成)

B.1 は音程ごとの暗意の強さを図式的に表したグラフである．縦軸が暗意の強さ，横軸が先に聴こえた音程を表す．青い線が似た性質を暗意する強さ，黄色い線が遠い性質を暗意する強さを表す．似た性質は，先の音程が1半音の時にもっとも強く暗意される．一方遠い性質は1半音では弱い暗意である．暗意の強さは，先の音程が6と7の間で逆転する．

終端記号の出力確率の事前分布には，次のベータ分布を用いた．

$$Beta(\lambda|\iota) = \frac{\lambda^{\iota_x-1}(1-\lambda)^{\iota_y-1}}{B(\iota)} \quad (B.3)$$

$$B(\iota) = \int_0^1 t^{\iota_x-1}(1-t)^{\iota_y-1} dt \quad (B.4)$$

λ の期待値が $\iota = (\iota_x, \iota_y)$ によって変化する．学習に用いたハイパーパラメータは $\iota_{I[n]} = (2, 2)$ とした．

サンプリング時のパラメータ

ギブスサンプリングのサンプル回数は5000回とした．最初の1000サンプルはburn in期間として破棄した．独立とみなせるサンプルを得るために80回ごとにサンプルを保存し，他の79サンプルは破棄する．

付録C 学習結果

本章で省略した学習結果を示す
 全ての非終端記号における $\theta_{I[N]}$ のグラフを以下に示す.

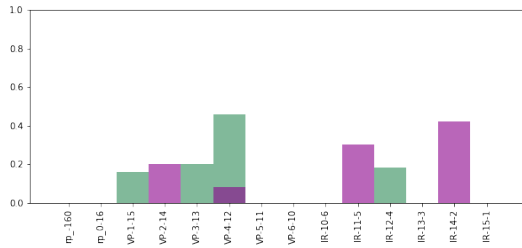


図 C.1: $\theta_{I[-16]}$

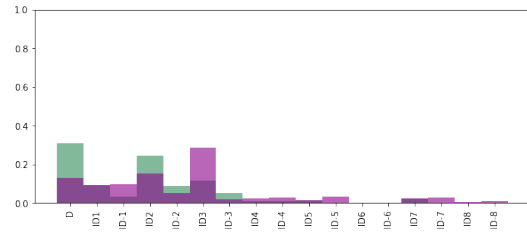


図 C.2: $\theta_{I[0]}$

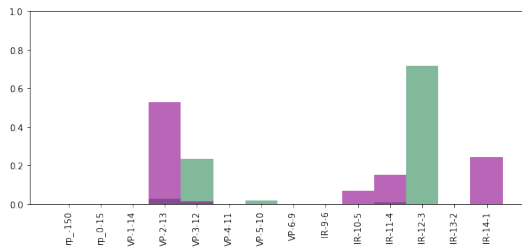


図 C.3: $\theta_{I[-15]}$

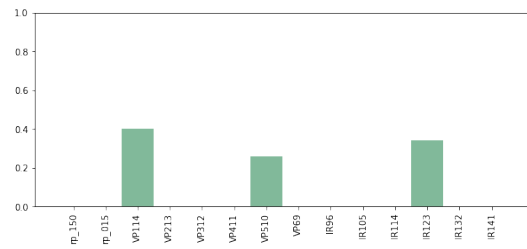


図 C.4: $\theta_{I[15]}$

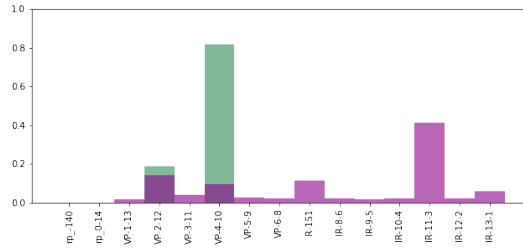


图 C.5: $\theta_{I[-14]}$

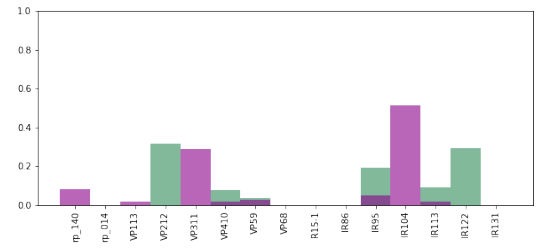


图 C.6: $\theta_{I[14]}$

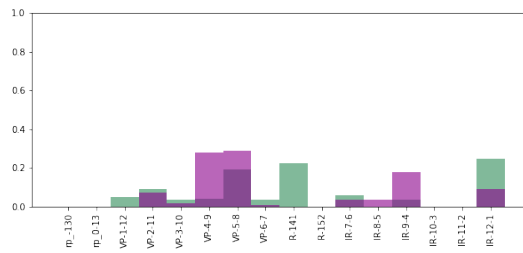


图 C.7: $\theta_{I[-13]}$

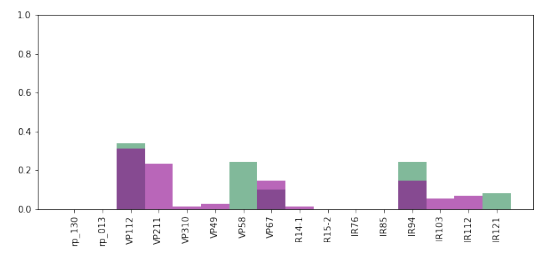


图 C.8: $\theta_{I[13]}$

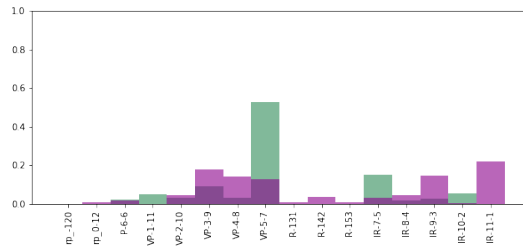


图 C.9: $\theta_{I[-12]}$

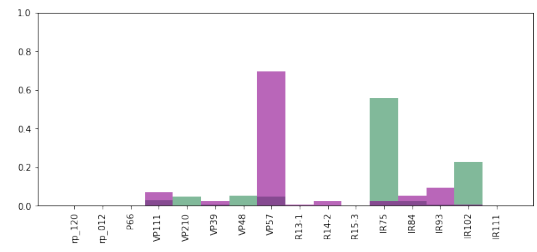


图 C.10: $\theta_{I[12]}$

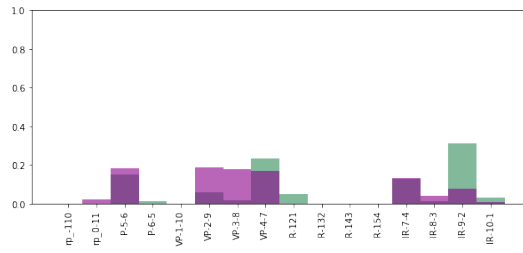


图 C.11: $\theta_{I[-11]}$

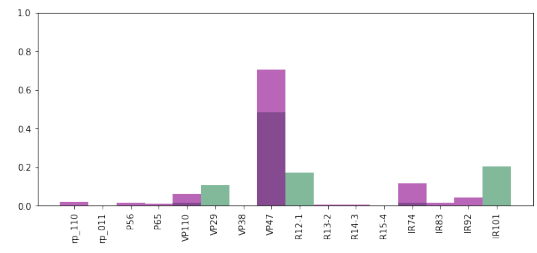
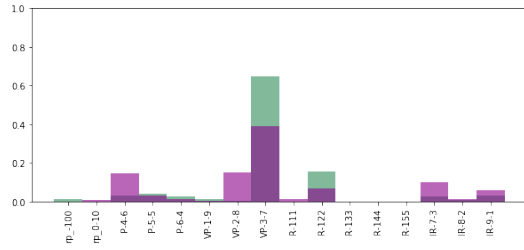
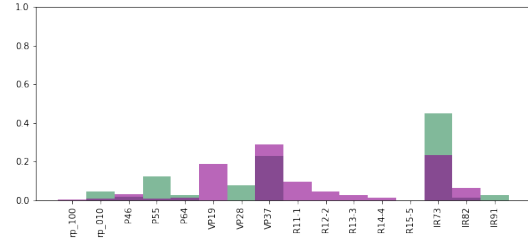


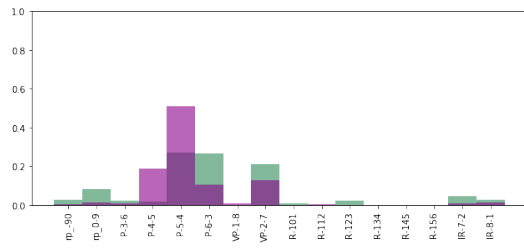
图 C.12: $\theta_{I[11]}$



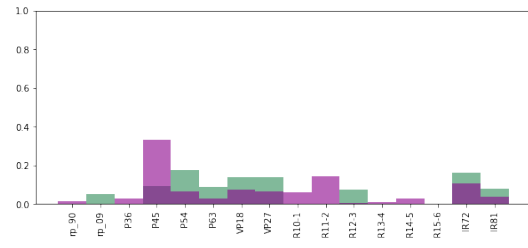
☒ C.13: $\theta_{I[-10]}$



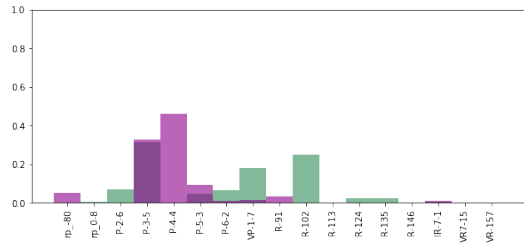
☒ C.14: $\theta_{I[10]}$



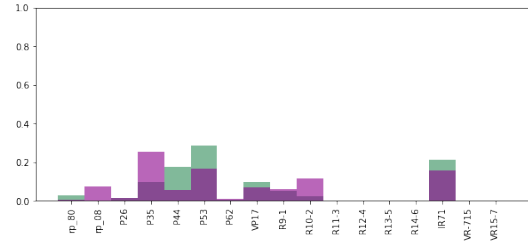
☒ C.15: $\theta_{I[-9]}$



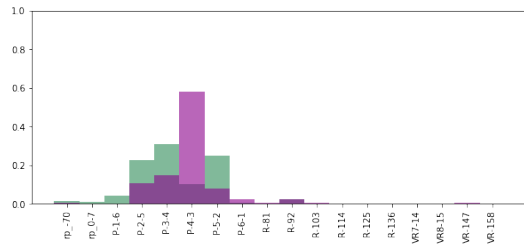
☒ C.16: $\theta_{I[9]}$



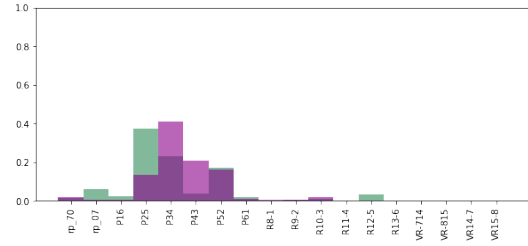
☒ C.17: $\theta_{I[-8]}$



☒ C.18: $\theta_{I[8]}$



☒ C.19: $\theta_{I[-7]}$



☒ C.20: $\theta_{I[7]}$

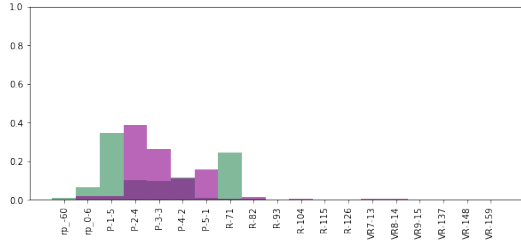


图 C.21: $\theta_{I[1-6]}$

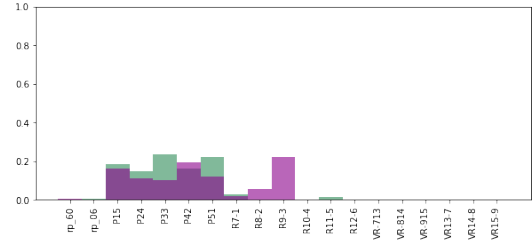


图 C.22: $\theta_{I[6]}$

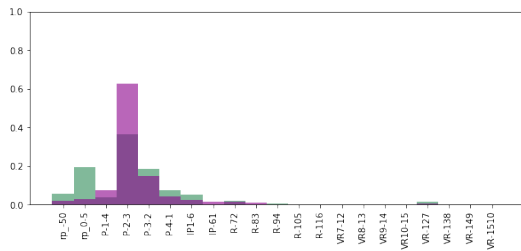


图 C.23: $\theta_{I[-5]}$

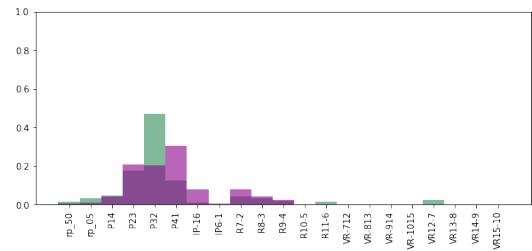


图 C.24: $\theta_{I[5]}$

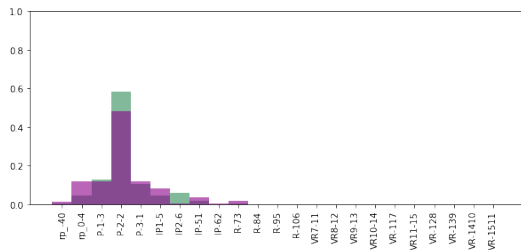


图 C.25: $\theta_{I[-4]}$

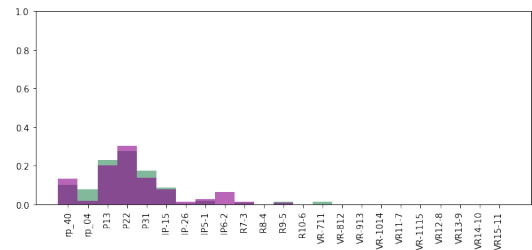


图 C.26: $\theta_{I[4]}$

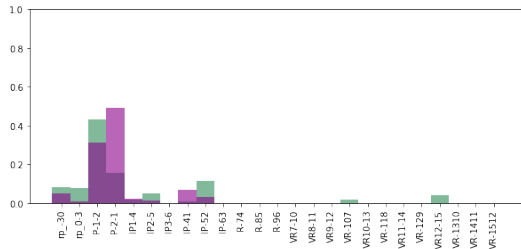


图 C.27: $\theta_{I[-3]}$

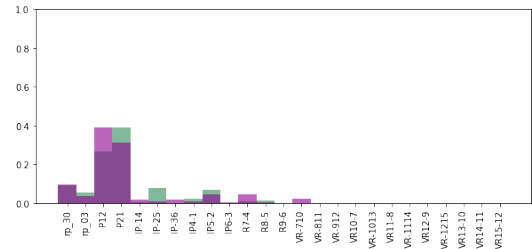


图 C.28: $\theta_{I[3]}$

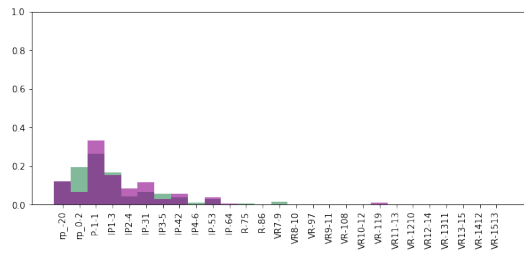


图 C.29: $\theta_{I[-2]}$

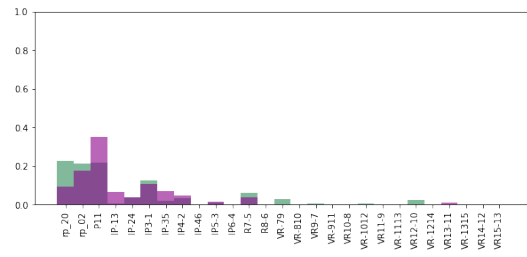


图 C.30: $\theta_{I[2]}$

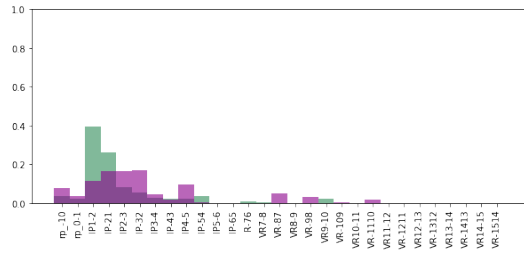


图 C.31: $\theta_{I[-1]}$

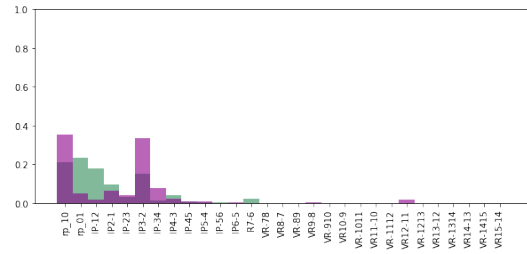


图 C.32: $\theta_{I[1]}$