

ベイジアンネットワークを用いた四声体和声付け: コードノードの有無による出力結果の比較

鈴木 峻平[†] 北原 鉄朗[‡]

[†] 日本大学大学院総合基礎科学研究科地球情報数理科学専攻

[‡] 日本大学文理学部情報科学科

1. はじめに

四声体和声は、伴奏がついた演奏で一番基本的なものの1つであり、これを自動生成する手法については既に様々な研究が行われている^{1)~3)}。既存研究では、コードシンボルをモデルに導入しているが、コードシンボルはその分類方法に曖昧性を含んでいる。例えば、コード C-E-G、E-G-C、G-C-E は同じコードシンボルである“C”で表される。“G-A-C-E”もまたテンションノートは時に省略されることもあるので“C”と表されるかもしれない。このようなコードシンボルをモデルが学習する場合、モデルで転回形やテンションノートは区別されない。一方、これらを正確に区別する場合は要素数が膨大になりデータスパースネスが起こる。そこで、以前我々が提案したベイジアンネットワークモデル⁴⁾を用いて、以下の仮説を客観評価や主観評価を行い検証する。

- (1) 転回形やテンションノートを区別しないモデルでは、適切に転回形やテンションノートが選ばれず、生成された和声は滑らかさに欠ける。
- (2) 転回形やテンションノートを区別したモデルでは、データスパースネスの為に限られた学習データでは適切な和声は生成されない。
- (3) コードシンボルを省略したモデルでは、コードシンボルを用いた結果よりも滑らかな和声が生成される。

2. 対象とする四声体和声生成手法

本研究の手法は基本的には先行研究⁴⁾と同様である。

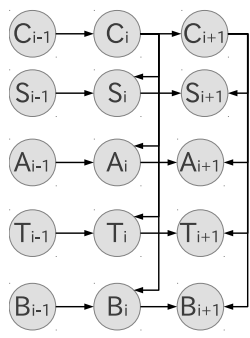


図1 コードモデル

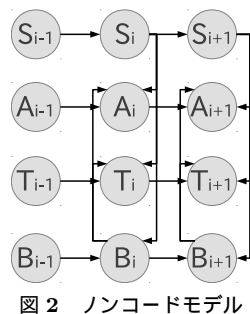


図2 ノンコードモデル

本研究では先行研究で設計したコードノードを含むモデル(コードモデル)とコードノードを含まないモデル(ノンコードモデル)を用いる。加えてこれまで用いていた仮説(1)を検証するコードモデルとは別に、仮説(2)を検証するために

転回形やテンションノートを分類したコードモデルを新たに作成した。そこで、仮説(1)のコードモデルである24種類のコードシンボルを学習したコードモデルをコードモデル1とし、仮説(2)で用いる110種類のコードシンボルを学習したコードモデルをコードモデル2とする。110種類は実際の学習データに現れたコードシンボルである。

3. 実験

3.1 讃美歌データベースの作成

賛美歌集⁵⁾から254曲抽出し、八長調に移調して使用した。これらのデータにはコードの情報が付与されていないので、「Band in a Box」⁶⁾のコード推定機能(精度95%以上)で推定した結果を付与した。

3.2 実験条件

我々は、四声体和声として学習するために全ての曲の8分音符よりも長い音符を8分音符に分割し、条件付確率表を学習した。この分割によって本来よりも同音連続の確率が高くなってしまったため、この確率を0.4倍する。

3.3 実験方法

和声の教科書⁷⁾に課題として載っている32曲のソプラノ課題を用いる。評価方法は、客観評価と主観評価を用いて行う。客観評価では以下に示す6つの基準で評価する。

- 項目1 不協和音を含む和音の総数
- 項目2 ノンダイアトニックノートの総数
- 項目3 最後のコードがCかどうか
- 項目4 3声部以上で同じ音名の音で構成される和音の総数
- 項目5 バスで連続する跳躍(完全4度以上)の総数
- 項目6 各声部内での音名の総数

主観評価は、作曲経験者3人に以下の基準で5段階評価を行ってもらった。(5:非常に良い、1:非常に悪い)

- 項目1 不協和音の少なさ
- 項目2 各声部の滑らかさ

以上の項目で評価した結果を比較した。

3.4 結果と考察

表1 客観評価 (C1:コードモデル1 C2:コードモデル2 N:ノンコードモデル)

項目	C1	C2	N	
項目1	1%	42%	8%	
項目2	0.07%	0%	0.4%	
項目3	100%	94%	100%	
項目4	38%	28%	20%	
項目5	16%	3%	12%	
項目6	alto	2.0	2.0	5.2
	tenor	2.6	1.6	4.9
	bass	3.3	2.1	4.8

Four-part Harmonization Using Bayesian Networks: Comparison of Outputs by Models With and Without Chord Nodes by Syunpei Suzuki, and Tetsuro Kitahara (Nihon University)
コード“C-E-G”はC, E, Gからなるコードを表す。

表 2 主観評価 (C1:コードモデル 1 C2:コードモデル 2 N:ノンコードモデル)

被験者	被験者 1			被験者 2			被験者 3		
	C1	C2	N	C1	C2	N	C1	C2	N
項目 1	4.81	3.28	4.58	2.53	1.38	2.91	4.75	1.56	3.91
項目 2	4.06	2.97	4.29	2.41	1.44	2.66	4.43	1.84	4.09

図 3 ノンコードモデルを用いた出力結果の一例

図 4 コードモデル 1 を用いた出力結果の一例

図 5 コードモデル 2 を用いた出力結果の一例

客観評価の結果を表 1、主観評価の結果を表 2 に示す。客観評価の項目 1~5 に関しては該当する和音の割合を示し、客観評価の項目 6 および客観評価は平均を示す。更に、図 3~5 に各モデルを用いた結果の一例を示す。

不協和音は、コードモデル 1 では全体の和音の 1%、ノンコードモデルでは全体の 8% が不協和音だった。コードモデル 1 ではコードと音符間の関係を直接学習したので、適切なコードが推論されれば適切なコードトーンが選ばれた。対照的に、コードモデル 2 では全体の和音の 42% が不協和音であった。主観評価では、同時性に関する基準に関してはノンコードモデルとコードモデル 1 の結果は平均 4.5 以上となった。コードモデル 2 では他と比べて低くなった。実際に図 4 ではコードが C, F, G の構成に対して図 3 では C, Dm, Em, F, G7, Am で構成されており、より多くのコードを使いつつ不協和音が少なかった。一方コードモデル 2 の結果(図 5)では、多くの不協和音が見られた。

コードモデル 1 ではノンダイアトニックノートは 1 つだけであった。対して、ノンコードモデルではいくつも見られた。コードモデル 2 ではひとつもなかったが、これはデータスパースネスにより結果のほとんどが同じコードとなったためだ

と考えられる。実際に図 3~5 では 1 つも現れなかった。

全ての曲でノンコードモデルとコードモデル 1 では最後のコードは C で終わっていた。これよりこれらのモデルは最後のコードがトニックコードで終わるといった基本的なルールを学習したと考えられる。対照的に、コードモデル 2 では C メジャーコードで終わっていない曲は 2 つだった。結果の一例(図 3~5)では 3 つともコード C で終わっている。

3 声部以上で同じ音名の音で構成される和音は、コードモデル 1 では全体の 38%、コードモデル 2 では 28% だった。これに比べ、ノンコードモデルでは 20% という結果になり、ノンコードモデルの結果よりもコードモデルの結果の方が多結果となった。実際にノンコードモデルの結果(図 3)では 3 つなのに対してコードモデル 1 の結果(図 4)では 7 つ、コードモデル 2(図 5)の結果では 5 つだった。

コードモデル 1 では、バスでの連続する大きな跳躍がノンコードモデルと比べ多い結果となった。コードモデル 2 ではコードモデル 1 よりも大きな跳躍は少なかった。実際にノンコードモデルの結果(図 3)では 1 つなのに対して、コードモデル 1 の結果(図 4)では 3 つと多い結果になった。

2 つのコードモデルでは、各声部の音名の総数は 2~4 だった。一方、ノンコードモデルでは 4 以上となった。実際に、コードモデルの結果でのアルトとテノールでは 1,2 種類の音が繰り返されているだけなので単調になっていた。しかし、ノンコードモデルでは各声部で滑らかなメロディになっていた。客観評価においても、2 人の被験者による連続性の評価はコードモデル 1 よりもノンコードモデルの方が少しではあるが高くなっていた。結果の一例でも、ノンコードモデルの結果ではアルト 4、テノール 5、バス 6 であるが、コードモデル 1 ではアルト 2、テノール 2、バス 3、コードモデル 2 の結果ではアルト 2、テノール 1、バス 2 というノンコードモデルよりもコードモデルの方が少ない結果となった。

4. 結 論

本論文では、以前我々が提案した四声体和声付けの手法を用いてコードシンボルの有無による結果の違いを検証した。コードシンボルには曖昧性があるため、3 つの仮定に基づいて 3 つのモデルの結果の比較を行った。結果は、ノンコードモデルの方がコードモデルよりもなめらかだった。今後は、コーパスを大きくし、禁則などの考慮をしていく。

参 考 文 献

- 1) Hild, H. et al. (1991). HARMONET: A neural net for harmonizing chorales in the style of J.S. Bach. *Advances in Neural Information Processing*, 4, 267-274.
- 2) Allan, M. et al. (2005). Harmonising chorales by probabilistic inference. *Advances in neural Information Processing Systems*, 17, 25-32
- 3) Liangrong, Y. et al. (2007). Automatic Generation of Four-part Harmony. *Conference on Uncertainty in Artificial Intelligence-Applications Workshop*, 268.
- 4) 鈴木峻平. et al. (2013). ペイジアンネットワークを用いた四声体和声付け:音の前後関係を考慮したモデルを用いた検討. 研究報告音楽情報科学 (MUS), 2013-MUS-99-9, 1-6.
- 5) 日本基督教団讃美歌委員会 (1982). 讃美歌・讃美歌第二編ともうたおう, 日本基督教団出版局.
- 6) Band-in-a-Box. (2009). PG music Inc.
- 7) 島岡 謙, et al., 和声理論と実習 (別巻), 音楽之社, 1967.