

形式概念に基づく Top-N 楽曲クラスタリングに関する一考察

Extracting Top-N Melody Clusters Based on Formal Concept Analysis

大久保 好章 原口 誠
Yoshiaki OKUBO Makoto HARAGUCHI

北海道大学大学院情報科学研究科コンピュータサイエンス専攻
Division of Computer Science, Graduate School of Information Science and Technology, Hokkaido University

In this paper, we discuss a method of finding *melody clusters* by *Top-N formal concept search*. Since a melody is a sequence of notes, it can be represented as a string on an alphabet Σ consisting of symbols denoting notes. We define a set of features as a set of substring extracted from a given melody. A melody md is said to have a feature f , if we can find a substring f' of md such that the edit distance between f and f' is less than or equal to a given threshold. Thus, by means of *approximate string matching*, we can obtain a formal context with respect to a given set of melodies. Under the formal context, our algorithm finds Top-N melody clusters each of which approximately shares some sequences of notes (features) as its intent.

1. はじめに

著者等はこれまで、文献 [原口 02] をその始まりとして、クリーク探索に基づくクラスタ抽出の研究を行ってきた (例えば [Haraguchi 06a, Haraguchi 06b, Okubo 05]). 対象とするデータ (個体) 群を、所与の類似関係のもとで無向グラフ表現し、その極大クリークを探索することで、クラスタ抽出を行なう。特に、評価値が上位 N (Top-N) のクラスタのみを、効率良くピンポイントに抽出することが大きな特徴である。

クラスタ抽出においては、クラスタの解釈・意味付けが重要であることが広く認識されており、文献 [Haraguchi 06b] では、こうした点を、形式概念解析 (Formal Concept Analysis) [Ganter 99] の枠組で議論している。形式概念解析では、外延 (個体集合) と内包 (共有属性集合) の組として概念を定義し、これを形式概念と呼ぶ。抽出すべきクラスタを形式概念に限定することで、クラスタ (外延) の解釈を、その内包の言葉で明確に語る事が可能となる。

文献 [Haraguchi 06b, Okubo 06] では、クラスタ抽出問題を、内包に関する制約を満たし、かつ、外延の評価値が Top-N の形式概念を求める問題として定式化し、その計算アルゴリズムを設計・実装した。それは著者らの従来アルゴリズムの拡張であり、そこでも分岐限定を利用した効率良い探索が可能である。

これまでの研究では文書データを対象に Top-N 形式概念クラスタの抽出を議論してきたが、その様々な分野への応用を展開するために、本稿では、当手法の楽曲データへの適用を試みる。具体的には、標準 MIDI ファイルから抽出されるメロディーのクラスタリングについて考察する。

クラスタを形成する個体、すなわち、メロディーは音符の時系列である。こうした時系列性を考慮し、ここでは、適当な長さの音符列を属性と捉え、メロディーがある音符列 (属性) を有するか否かを、近似文字列照合により判定する。その結果、いくつかの音符列を近似的に共有するメロディー集合が、ひとつのクラスタを形成する。

連絡先: 大久保 好章・原口 誠

北海道大学大学院情報科学研究科コンピュータサイエンス専攻

〒060-0814 札幌市北区北14条西9丁目

TEL : 011-706-7161 (FAX 兼用)

E-mail : { yoshiaki, mh }@ist.hokudai.ac.jp

メロディークラスタの評価にあたっては、ユーザが実際にそれを妥当なものと感じるか否かという、極めて認知的・感性的な判断が不可欠であるが、形式概念に基づくクラスタを抽出することで、こうした感性的な判断を、記号レベルで解釈できる可能性がある。

2. 準備

個体 (*individual*) の集合 \mathcal{O} 、および、属性 (*feature*) の集合 \mathcal{F} に対して、関係 $R \subseteq \mathcal{O} \times \mathcal{F}$ を考える。この時、タプル $\langle \mathcal{O}, \mathcal{F}, R \rangle$ を、形式文脈 (Formal Context) と呼ぶ。 $(o, f) \in R$ の時、個体 o は属性 f を有すると言う。個体 o が有する属性の集合 $\{f \in \mathcal{F} \mid (o, f) \in R\}$ を、 $F(o)$ で参照する。

形式文脈 $\langle \mathcal{O}, \mathcal{F}, R \rangle$ に関して、写像 $\varphi: 2^{\mathcal{O}} \rightarrow 2^{\mathcal{F}}$ および $\psi: 2^{\mathcal{F}} \rightarrow 2^{\mathcal{O}}$ を考える。ここで、個体集合 $O \subseteq \mathcal{O}$ と属性集合 $F \subseteq \mathcal{F}$ について、

$$\varphi(O) = \{f \in \mathcal{F} \mid \forall o \in O \ f \in F(o)\} = \bigcap_{o \in O} F(o),$$

$$\psi(F) = \{o \in \mathcal{O} \mid F \subseteq F(o)\}$$

とする。つまり、 φ は O 中のすべての個体が共有する属性の集合を、一方、 ψ は F 中のすべての属性を有する個体の集合を返す写像である。

これら写像のもと、個体集合 $O \subseteq \mathcal{O}$ と属性集合 $F \subseteq \mathcal{F}$ について、 $\varphi(O) = F$ かつ $\psi(F) = O$ が成り立つ時、 O と F の組 $FC = (O, F)$ を形式概念 (Formal Concept) と定める。ここで、 O と F をそれぞれ FC の外延 (*extent*)、および、内包 (*intent*) と呼ぶ。 φ と ψ の定義より、 $\psi(\varphi(O)) = O$ かつ $\varphi(\psi(F)) = F$ であることは明らかである。すなわち、形式概念とは、写像 φ と ψ に関して閉じた (*closed*) 個体集合 O と属性集合 F の組で与えられる。 O は、 F 中のすべての属性を有する個体のみから成り、かつ、それら以外にこうした個体は存在しない。同様に、 F は、 O 中のすべての個体に含まれる (共有される) 属性のみから成り、かつ、それら以外にこうした属性は存在しない。以降では、閉包 (*closure*) を生成する場合関数 $\varphi \circ \psi$ および $\psi \circ \varphi$ をどちらも *closure* で表す。

3. 標準 MIDI ファイルからのメロディーデータ抽出

MIDI (Musical Instruments Digital Interface) は、シンセサイザーに代表される電子楽器 (機器) やパソコン等の間でデータをやり取りするための通信プロトコルである。MIDI 形式の音楽データを扱う際のファイル形式は、標準 MIDI ファイル (SMF : Standard MIDI File) として標準化されており、ユーザ間で音楽データをやり取りする際には、これが手軽に広く用いられている。本稿においてもこうした SMF 形式の音楽データを扱うものとする。

SMF には、各楽器 (チャンネルと呼ぶ) 毎に、それに対する演奏情報がデルタタイムとイベントの組の時系列として記述される。時系列を構成する組は『いつ (when), 何 (what) をするか』を正確に定めたものであり、これら命令に従ってチャンネルが制御される。

イベントには、MIDI イベント・SysEx イベント・メタイベントの 3 種類が定められており、曲のメロディーに関する情報は、MIDI イベントに記述される。具体的にメロディーを抽出する際には、MIDI イベントの発音メッセージ (ノートオン) および消音メッセージ (ノートオフ) に注目すればよい。前者は鍵盤を押す操作、後者は鍵盤を離す操作に対応し、これらの情報から、『どの音をどのくらいの長さ鳴らし続けられればよいか』がわかる。すなわち、楽譜に記載された音符列に相当する情報が抽出できる。

より形式的に述べると、各 SMF から、音名 p_i とその持続時間 l_i の組 (p_i, l_i) の時系列

$$\langle (p_1, l_1), \dots, (p_i, l_i), \dots, (p_M, l_M) \rangle$$

をメロディーとして抽出する。ここで、 (p_i, l_i) を音符と呼ぶ。

SMF の詳細については、文献 [AMEI 98] 等を参考にされたい。

4. メロディーデータに関する形式文脈の生成

標準 MIDI ファイルの集合を SMF とする。各 $SMF_i \in SMF$ から抽出されるメロディーを

$$melody(SMF_i) = \langle (p_{i1}, l_{i1}), \dots, (p_{iM_i}, l_{iM_i}) \rangle$$

とした時、 $MD = \{melody(SMF_i) \mid SMF_i \in SMF\}$ を SMF に関するメロディーデータベースと呼ぶ。本稿では、 MD 中のメロディーを対象に、形式概念としてのメロディークラスタの抽出を試みる。

形式概念解析の枠組でクラスタ抽出を行なう際には、各メロディーを、それが有する属性の集合として表現する必要がある。メロディーを考える上で、音符の時系列性を考慮することは不可欠であるため、ここでは、音符列をひとつの属性と考えることにする。すなわち、属性 f_i は

$$\langle (p_{i1}, l_{i1}), \dots, (p_{ik}, l_{ik}) \rangle$$

なる長さ k の音符列である。

先に定義した通り、形式概念の外延は、そこに含まれる個体が内包中の属性をすべて共有していることを意味する。しかし、個体としてメロディーを考える場合、異なるメロディーが、いくつかの音符列を厳密に共有することは非常に稀であると思われる。つまり、複数のメロディーが形式概念の外延を形成することは、ほとんど期待できないであろう。そこで本稿で

は、メロディーがある属性を有するか否かの判断を、文字列のアラインメントを考えることで近似的に行なう。

いま、任意の音符 (p_1, l_1) を記号とするアルファベット Σ を考える。すると、メロディーおよび属性はすべて、 Σ 上の文字列と考えることができる。ここで、メロディーをターゲットテキスト T 、属性をパターン P と考えて、近似文字列照合を行なう。その結果、 T が、 P と編集距離が d 以内の部分文字列を含むことが分かれば、メロディー T は属性 (音符列) P を有すると考える。

Remark:

属性の有無を近似的に判定する方法は他にもある。例えば、音符 (p_i, l_i) のうち、音名 p_i のみに注目し、持続時間 l_i を無視した上で属性の有無を考えることも可能である。逆に、持続時間のみ注目し、音名を無視してもよい。こうした抽象化を行なった上で、上記のアラインメントを考えることももちろん可能である。

言うまでもなく、任意の音符列 (Σ 中の文字列) が、近似文字列照合のパターン (属性) になり得るが、長過ぎるパターンや短過ぎるパターンは属性として適当ではない。また、ランダムに並べた音符列にも音楽的な意味は期待できない。よってここでは、既存のメロディーに実際に含まれる音符列を属性として用いることにする。具体的には、メロディー $\langle (p_1, l_1), \dots, (p_M, l_M) \rangle$ から、長さ k の連続した任意の音符列

$$\langle (p_i, l_i), (p_{i+1}, l_{i+1}), \dots, (p_{i+k-1}, l_{i+k-1}) \rangle$$

を属性として抽出する ($1 \leq i \leq M - k + 1$)。

いま、所与の既存メロディー (の集合) から抽出された属性 (音符列) の集合を \mathcal{F} 、編集距離の閾値を d とする。この時、

$$R = \{(md, f) \mid md \in MD, f \in \mathcal{F} \text{ かつ } md \text{ は } f \text{ と編集距離が } d \text{ 以内の部分文字列を含む.}\}$$

を考えると、

$$\langle MD, \mathcal{F}, R \rangle$$

は、各メロディーに関する属性の有無を近似的に判断した場合の形式文脈を与える。

5. Top-N 楽曲メロディークラスタの抽出

ここでは、形式文脈 $\langle MD, \mathcal{F}, R \rangle$ のもとで、少なくとも δ の音符列 (属性) を含む内包を有する形式概念のうち、外延に含まれるメロディー数が上位 N であるものを Top-N クラスタとして抽出する。その計算には、既存アルゴリズム [Haraguchi 06b, Okubo 06] が利用可能である。

まず最初に、無向グラフ $G = (MD, E)$ を作成する。ここで、辺集合

$$E = \{(md_i, md_j) \mid md_i, md_j \in MD (i \neq j) \text{ かつ } |F(md_i) \cap F(md_j)| \geq \delta\}$$

とする。すなわち、各メロディーペアについて、それらが δ 以上の音符列を共有する場合に限り、それらの間に辺を作成する。こうして作成されたグラフにおいて、内包制約を満たすメロディークラスタは、クリークを形成する。よって、 G のクリークを探索することで Top-N クラスタの抽出が可能となる。

Top-N クラスタの抽出アルゴリズムは、分枝限定深さ優先クリーク探索 [Tomita 03] の拡張である。そこでは、内包制約

を満たさない形式概念や、外延サイズが Top- N に成り得ない形式概念の探索が積極的に枝刈りされ、効率良く Top- N クラスタが抽出される。なお、アルゴリズムの詳細については、文献 [Haraguchi 06b, Okubo 06] を参照されたい。

6. おわりに

本稿では、形式概念に基づく Top- N クラスタ抽出手法により、楽曲 MIDI データからメロディークラスタを抽出する枠組について考察した。メロディーを構成する音符の時系列性を考慮し、ここでは、適当な長さの音符列を属性と考える。メロディーや属性は、音符を記号とする文字列で表現可能なことから、メロディーがある属性を有するか否かを、近似文字列照合により判定する。その結果、いくつかの属性を近似的に共有するメロディーの集合が、ひとつのクラスタを形成する。

現在、本手法によるメロディークラスタ抽出実験を行なっている。抽出されるメロディークラスタの実際の評価については、稿を改めて報告したい。

ここでの基本アイデアは、メロディーデータのみならず、様々なドメインの時系列データに対して議論できるものであり、それらについての考察も大変興味深いと考えている。

参考文献

- [Ganter 99] B. Ganter and R. Wille: Formal Concept Analysis - Mathematical Foundations, Springer, 1999.
- [AMEI 98] (社) 音楽電子事業協会 (AMEI) : MIDI 1.0 規格書 (日本語版 98.1), リットーミュージック, 1998.
- [Tomita 03] E. Tomita and T. Seki: An Efficient Branch-and-Bound Algorithm for Finding a Maximum Clique, Proceedings of the 4th International Conference on Discrete Mathematics and Theoretical Computer Science - DMTCS'03, Springer-LNCS 2731, pp. 278 - 289, 2003.
- [原口 02] 原口 誠: 最適クリーク探索に基づくデータからの概念学習, 人工知能学会研究会資料, SIG-FAI-A202, pp. 63 - 66, 2002.
- [Haraguchi 06a] M. Haraguchi and Y. Okubo: A Method for Pinpoint Clustering of Web Pages with Pseudo-Clique Search, Federation over the Web, International Workshop, Dagstuhl Castle, Germany, May 1 - 6, 2005, Revised Selected Papers, Springer-LNAI 3847, pp. 59 - 78, 2006.
- [Haraguchi 06b] M. Haraguchi and Y. Okubo: An Extended Branch-and-Bound Search Algorithm for Finding Top- N Formal Concepts of Documents, Proceedings of the 4th Workshop on Learning with Logics and Logics for Learning - LLLL'06, pp. 41 - 47, 2006.
- [Okubo 06] Y. Okubo and M. Haraguchi: Finding Conceptual Document Clusters with Improved Top- N Formal Concept Search, Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence - WI'06, pp. 347 - 351, 2006.

- [Okubo 05] Y. Okubo and M. Haraguchi: Finding Significant Web Pages with Lower Ranks by Pseudo-Clique Search, Proceedings of the 8th International Conference on Discovery Science - DS'05, Springer-LNAI 3735, pp. 346 - 353, 2005.