

階層的文化伝達モデルに基づく芸術データからの知識参照ネットワーク推定

京大情報^A
中村栄太^A

Estimation of knowledge reference network from art data based on hierarchical cultural transmission model

^AGrad. School of Info., Kyoto Univ.
E. Nakamura^A

学術文献科学や芸術データ分析などの文化進化研究では、知識の参照構造を表すネットワークから経験則を抽出する方法が用いられている。しかし、文化現象ではこのような知識参照ネットワークの直接観測は難しい状況も多い。本研究では、音楽のコード進行データを具体的な対象として、創作物データの進化モデルに基づく統計推論の枠組みを用いて、創作物データからその背後にある知識参照ネットワークを推定する方法を調べる。進化過程では、各創作物は過去の創作物を参照して、その要素に変異が加わることで生成されると考える。また、楽曲は複数のコード進行の断片から構成され、その各々が異なる楽曲を参照し得るため、この階層的な参照構造を含む文化伝達モデルを確率過程として定式化する。

各曲 n は記号列の断片 w_{nm} の集合として表し、断片 $w_{nm} = (w_{nm\ell})_{\ell=1}^L$ の長さ L は固定して考える ($1 \leq m \leq M_n$, $L = 3$)。各 $w_{nm\ell}$ はコード記号を表す。また曲の制作時刻を t_n とする。各曲の断片 w_{nm} は、既存の曲 n' の断片 $w_{n'm'}$ を参照して、変異確率 $P_{\text{mut}}(w_{nm}|w_{n'm'}) = \prod_{\ell} \phi_{w_{n'm'\ell}}(w_{nm\ell})$ により生成されるとする。ここで、 $\phi_{a'}(a)$ はコード記号 a' から a への置換確率を表す。既存データの中で参照曲 n' が選ばれる選択確率は、曲の制作時刻が近いほど高くなり (近親性)、また曲に固有の影響度 (適応度) $v_{n'}$ に比例して決まるとする: $\pi_{n'}^{(t_n)} \propto e^{-(t_n-t_{n'})/\tau} e^{v_{n'}}$ 。ここで、参照曲の数は、その候補の数に比べて少数であるという現実的な状況を表すため、各曲の選択確率 $\pi^{(n)} = (\pi_{n'}^{(n)})$ は小さな集中度 α を持つ Dirichlet 分布により生成されると仮定する。また、断片 m' が選ばれる確率 $P_{\text{sel}}(m'|n')$ は、断片ごとの適応度 $f_{n'm'}$ に比例して決まるとする: $P_{\text{sel}}(m'|n') \propto e^{f_{n'm'}}$ 。この過程は、以下の式で表される。

$$P(\{w_{nm}\}_m) = \int d\pi^{(n)} \text{Dir}(\pi^{(n)}; \alpha \bar{\pi}^{(t_n)}) \prod_m \left[\sum_{n'} \pi_{n'}^{(n)} \sum_{m'} P_{\text{sel}}(m'|n') P_{\text{mut}}(w_{nm}|w_{n'm'}) \right]$$

本研究では、上記モデルをバイズ拡張した統計モデル (図 1A) を用いて、与えられたデータ $\{w_{nm}, t_n\}$ からモデルパラメーターと参照ネットワークに対応する事後確率 $\tilde{\pi}_{n'}^{(n)} = \sum_{m,m'} P(n', m'|n, m)/M_n$ をマルコフ連鎖モンテカルロ法により推定する手法を導出した。合成データを用いた解析により、曲ごとの参照曲数が小さい場合は、参照ネットワークの推定精度がバイズモデルにより向上することが確かめられた (図 1B,C)。また、推定精度は、記号の多様性に関する条件により、変異率の増加関数とも減少関数ともなり得ることが示された。現実の音楽データを用いた分析では、和声機能を保つ置換確率が学習され、変異率の大きな時代変化や音楽ジャンルを反映したネットワークのコミュニティ構造の存在などが明らかになった。本手法は、作者-作品などの階層的構造に対してや連続量で特徴付けられる創作物データにも拡張でき、学術文献や絵画データの分析にも有用であると考えられる。

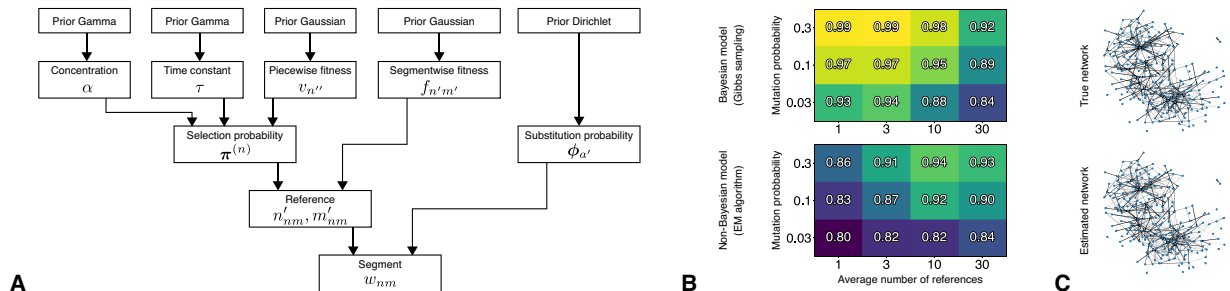


図 1 A: 進化モデルの変数。B: 合成データでの参照量 (入次数) の推定精度。C: 推定ネットワークの例。