

E-074

可変長Nグラムモデルを用いたジュウシマツの歌構造の解析に関する研究

Syntactic analysis of Bengalese finch's song using Variable length N-gram model

常田 宏和[†]
Hirokazu Tokida笹原 和俊[†]
Kazutoshi Sasahara西野 哲朗^{†‡}
Tetsuro Nishino高橋美樹[†]
Miki Takahasi岡ノ谷一夫[†]
Kazuo Okanoya

1 はじめに

鳥が歌と呼ばれる音声を学習する過程と、人間が音声言語を学習する過程の間には多くの共通点があり、鳥の歌が人間の音声言語獲得過程を解明するためのモデルになるのではないかと考えられている [1].

本研究では、ジュウシマツという鳥の歌を解析の対象とした。ジュウシマツの歌は音素と呼ばれる歌の最小構成要素と、毎回特定のパターンで現れる音素の組合わせであるチャンクで構成されている。ジュウシマツの歌は、複数のチャンクや音素が繰り返したり並び替えられたりして構成されており、固定的なパターンが多い他の鳥の歌と比べ複雑であることが分かっている。

歌構造を解析する手法として、可変長Nグラムモデルを用いた手法が提案されているが [3], 歌い間違い等のノイズが歌に含まれてしまうことや、その解析精度の問題がある。そこで本研究では、可変長Nグラムモデルを改良した解析手法について提案し、検証を行う。

2 可変長Nグラムモデル

単語予測に関する手法で、直前の $(N-1)$ 個の単語から次に現れる単語を予測するNグラムモデルという手法があるが、このNの値を単語の現れる履歴によって動的に変化させるモデルを可変長Nグラムモデルと呼ぶ [2].

可変長Nグラムモデルは予測接尾木を用いて構築される。予測接尾木では各枝には単語が、各節点には互いに異なった単語列がラベル付けされている。子節点のラベルは、親節点のラベルに、親節点からの枝のラベルを足したものになる。また、各節点には単語の遷移確率が割り当てられる。この予測接尾木は最初は根のみが存在し、

確率分布の距離に重み付けを行った関数 $\Delta H(\omega\alpha, \alpha)$ があらかじめ設定した閾値を越えた節点を、新たな節点として追加していくことにより構築する。

$$\Delta H(\omega\alpha, \alpha) = P(\omega\alpha) \times \sum_{\omega'} P(\omega'|\omega\alpha) \log \frac{P(\omega'|\omega\alpha)}{P(\omega'|\alpha)} \quad (1)$$

歌の解析においては、ジュウシマツの音素を可変長Nグラムモデルの単語に対応付ける。また、予測接尾木の葉節点を有限オートマトンの各状態に対応させて、視覚的に把握しやすい有限オートマトンに変換する。

3 提案手法

3.1 提案手法の流れ

まず前節で述べた関数 ΔH の閾値を決定し、鳥の歌の記号列データから可変長Nグラムの予測接尾木を作りこれをオートマトンに変換する。このオートマトンからノイズを除去し、状態のマージを行う。

3.2 ノイズの除去

各々の状態における単語の遷移確率を計算し、一定値以下の遷移はノイズとして除去する。この時の閾値をカットオフレベルと呼ぶ。またノイズ除去の結果、その状態への遷移がない状態やその状態からの遷移がない状態が現れた場合、これらも同様にノイズと判断し除去する。

3.3 状態のマージ

可変長Nグラムモデルをより簡潔に表わすために、次の条件を満たす2つの状態が存在する時、これらの状態を1つにまとめることにする。これを状態のマージと呼ぶ。

- ある状態に入る遷移がひとつだけであり、かつその直前の状態から出る遷移がひとつである時
- 状態のラベルが等しく、直後の状態も等しい二つの状態が存在する時

[†]電気通信大学電気通信学部情報通信工学専攻

[‡]理化学研究所脳科学総合センター生物言語研究チーム

^{††}電気通信大学電気通信学部情報工学科

3.4 可変長Nグラムモデルにおける閾値の決定

可変長Nグラムモデルにおいて、妥当性の高いモデルを構築するためには、予測接尾木を構築する際に扱う関数 ΔH の閾値を適切に設定しなければならない。また、個体ごとに適切な閾値は変動するため、一定の値が必ずしも正しいとは言えない。そこで視覚的に構造が把握しやすく、かつ冗長性が最も低いと考えられる状態数が最小の時の閾値を解析に用いることにする。しかし状態数が少ない場合、必要な状態を除去してしまい、与えたデータが十分に表わしていない可能性がある。そこで解析結果の検証を行い、用いた閾値が適切であるか確認を行った。

4 解析結果

提案手法を用いて、鳥の歌構造の解析を行い、解析の精度についての検証を行った。解析の精度は、可変長Nグラムモデルが与えられた記号列を不足なく表わしているかを確認することにより行った。

次の記号列はジュウシマツの歌の記号列データの一例である。

```
aaaaacbcddefghfghijfkfdefghijkfdefghijfghijfghijfghij
ijkfefbcbcddefghijkfdefghijkfbcbcddefghijkfdefghij
jkfdefghddefghijkfbcbcddefghijkfdefghijkfbcdbddef
fghijkfb
```

この歌の記号列データを提案手法で解析した結果が図1である。この時のノイズのカットオフレベルは10%、 ΔH の閾値は3.4節の結果に従い0.01とした。

解析結果を記号列データと比較してみると、与えたデータ中97.5%は表現できていることが確認できた。一例表現できないデータが存在したが、閾値を極端に下げない限り表現することができなかつたため、このデータは非常に稀なノイズを含んでいたと考えてよい。また、efやghi, jkfなどの特徴的なチャンクをうまく表現できている。さらに、gbdなどといった出現確率の低いノイズと考えられる記号列を除去できていることも確認できる。これらのことから、提案手法による解析はノイズを除去し、精度のよい解析を行えたといえる。

しかし、不完全な点もある。記号列の最初に存在するefghefという記号列は出現確率が低く、ノイズとして除去されてしまっている。しかしこの記号列は、出現位置に規則性のないノイズと違い、毎回歌の初期に現れているため、ノイズではない可能性が高い。このように統計的な手法では、出現確率は低いものの特定の位置に出現

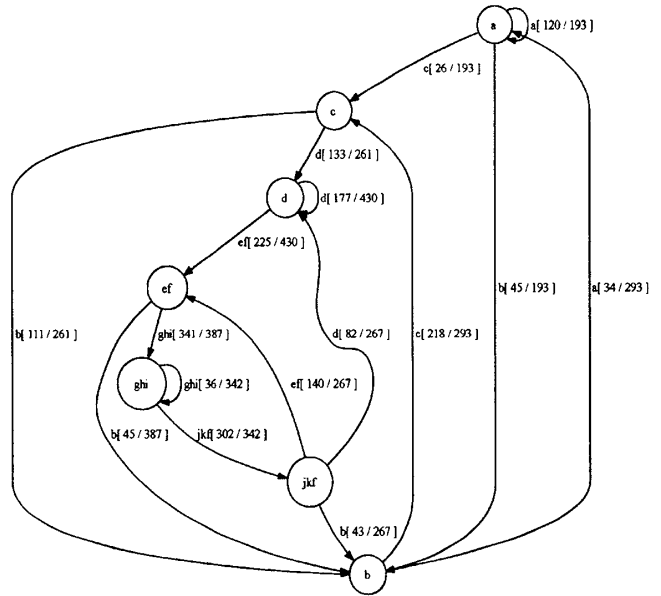


図1: ジュウシマツの歌の提案手法による解析結果

するようなチャンクをうまく表現できないという問題があり、改良の余地があるといえる。

5 おわりに

本研究では、可変長Nグラムモデルを改良したジュウシマツの歌構造の解析手法を提案し、その検証を行った。その結果、提案手法によってノイズを除去し、適切な閾値を用いて精度のよい歌構造の解析が行えることが確認できた。

しかし、特定の位置に現れるが出現確率は低いチャンクをうまく表現できないという問題がある。また、関数 ΔH の閾値の決定方法についても、稀に他の閾値を用いたの方が妥当性の高い解析が行えることがあるなど、問題がある。これらの問題を解決することが、今後の課題である。

これらの課題が解決できれば、本手法によって発達のデータや他個体間の比較を容易に行うことができると考えられる。

参考文献

- [1] 岡ノ谷一夫: “小鳥の歌からヒトの言葉へ”, 岩波書店, 2003
- [2] 北研二: “確率的言語モデル”, 東京大学出版会, 1999
- [3] 河村奨: “可変長Nグラムモデルの行動解析への応用”, 千葉大学卒業論文, 2001