

ノンパラメトリックベイズ二重分節解析器による 時系列データの分節化と生成

Segmentation and Generation of Time-series data
Using Nonparametric Bayesian Double Articulation Analyzer

谷口 忠大*¹ 古林 邦彬*¹ 長坂 翔吾*¹
Taniguchi Tadahiro Kobayashi Kuniaki Nagasaka Shogo

*¹ 立命館大学

The Japanese Society for Artificial Intelligence

In this paper, we describe a nonparametric Bayesian double articulation analyzer for modeling human motions and driving behaviors. Double articulation structure is a two-layered hierarchical structure underlying in spoken language. A double articulation analyzer estimates the hidden structure of observed data by segmenting and chunking it. We developed a double articulation analyzer by using a sticky hierarchical Dirichlet process HMM (sticky HDP-HMM), a nonparametric Bayesian model, and an unsupervised morphological analysis based on nested Pitman-Yor language model which can chunk given documents without any dictionaries. We describe unsupervised segmentation and generation approach to implicitly structured human behavioral data.

1. はじめに

連続時間の実空間上で類似した特徴量が持続的に観測されることによりひとまとまりとなる音素と、その離散的な符号の定型的な系列としてまとまる単語の二層構造によって音声信号は捉えられる。音声言語とは、人間によって、そのような構造を持つように生成され、そのような構造を前提として認識されているデータと考えることができる。人間の言語的情報処理に関わる脳機能が、この構造を好んでいるということを示唆しているように思われる。この二段階の分節構造を二重分節構造と呼ぶ。

二重分節構造は定型的なパターンの組み合わせにより生成される時系列データが持つ構造であり、人間が計画し動作することにより生成される時系列データが持ちやすい構造ではないかと考えられる。

この着想から筆者らは二重分節構造を前提とし、対象時系列データから教師なし学習により、音声認識における、音響モデルに相当するモデルや、言語モデルに相当するモデルを自動的に獲得する手法について研究を行ない、動作時系列データや、自動車運転データに対する適用を行ってきた。この手法を二重分節解析器と呼ぶ。

二重分節解析器は時系列データから、音声認識における言語モデルや音響モデルにあたる知識を獲得する学習モジュールである。人間の幼児は言語獲得において、親の発話から自動的に単語を獲得しており、二重分節解析に相当する学習を行なっているように見える。言語獲得の計算論を構築する上でも二重分節解析器の構成の意義は深いだろう。本講演ではノンパラメトリックベイズ二重分節解析についての概説を行う*¹。

2. ノンパラメトリックベイズ二重分節解析

2.1 データ生成過程

図 1 に沿って、二重分節解析器で前提とするデータ生成過程について説明する。まず、二重分節構造においては、言語モデルが存在し、単語遷移により隠れ状態列として単語列が生成されていると考える。それぞれの単語は複数の隠れ状態の遷移から成り立っている。一つの隠れ状態は確率的に自己遷移を続け、ある程度の時間、その状態に滞在し、確率的に次の状態へと遷移する。状態がある隠れ状態にとどまっている間、それぞれの隠れ状態に対応する出力分布から連続なデータを出力する。

二重分節解析器の議論では、音声のみならず多様なデータを扱うため、音声認識における音素にあたるものを出力素、音響モデルに相当するものを出力素モデルと呼ぶことにする。言語モデルに相当する確率モデルは、そのまま、言語モデルと呼ぶ。隠れた二重分節構造を持つ時系列データは言語モデル、出力素モデル、及び隠れ状態の滞在時間を表現するモデルの三つの要素により二重分節構造を持つデータの生成過程が表現される。

2.2 近似的推定

二重分節解析器が解かねばならない問題は、観測された連続な時系列データのみから適切な言語モデルと出力素モデルを教師なしで推定するという問題である。この時に出力素の総数は未知であり、言語モデルに含まれる語彙数も未知である。このような、問題を旧来のパラメトリックな手法で行おうとすると、膨大な数のモデル選択問題と共に解かねばならず、現実的には困難となる。筆者らはノンパラメトリックベイズ理論で提案されている手法を応用することで近似的なマルコフ連鎖モンテカルロ法を用いて解くことを提案している。

時系列データから言語モデルと出力素モデルを同時推定しようとする時に、問題となるのは単語数も出力素数も未知であるということである。そのため理論的には潜在的に無限の単語を想定する必要がある。Mochihashi らによって提案された Nested Pitman-Yor language model (NPYLM) は未観測の単語を含めた無限個の単語に対して適切な確率を与えることの出来るノンパラメトリックベイズ理論に基づく言語モデルである [2]。その特性を用いて Mochihashi らは辞書なしで与えられたデー

連絡先: 谷口忠大, 立命館大学情報理工学部, 〒 525-8577 滋賀県草津市野路東 1-1-1, 077-561-5839, taniguchi@ci.ritsumeiji.ac.jp

*¹ 本稿の内容は人工知能学会誌特集号においてより詳しく記述しているので参照されたい [1]

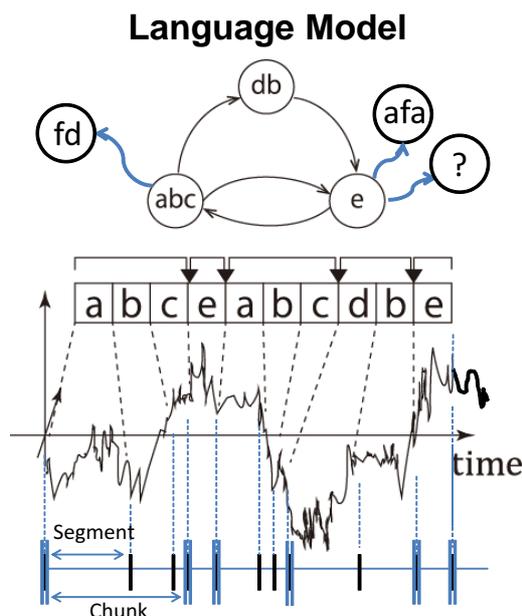


図 1: 二重分節解析のイメージ図

ただけから形態素解析を行なうベイズ教師なし形態素解析を提案している。二重分節解析器では NPYLM を単語列の生成を行なう言語モデルとする。単語列は時間方向に確率的な引き伸ばしを受け、各出力素の出力分布に従ってデータを出力する。

このような過程で生成されたデータからギブスサンプリングによって隠れ状態列の推定を行おうとすると、無限の語彙を含んだ言語モデルの確率を考慮しないといけなくなる。これは、計算量的に現実的に解くことが困難である。そこで、隠れ状態列をサンプリングする際に、複雑な言語モデルの影響を無視して、代わりにデータから学習したマルコフモデル (bigram 確率) を用いるという近似を [3] 採用している。この近似により、ギブスサンプリングはデータから隠れ状態列の無限隠れマルコフモデル [4] を用いた推定、得られた隠れ状態列からの NPYLM を用いた教師なし形態素解析という二段階の推定過程に分離している [3]。最終的なアルゴリズムは観測された時系列データに対して iHMM で符号化した後に、NPYLM で教師なし形態素解析 (チャンク化) するという、二段階の処理を繰り返し行なうというアルゴリズムになる。

3. 二重分節構造の推定と応用

3.1 全身動作データ

Taniguchi らは [3] において二重分節解析器 (Double Articulation Analyzer) を提案し、人間の動作データに適用した。ロボットによる教師なしの模倣学習においては、ロボットは学習すべき単位動作を自ら発見する必要がある。このタスクにおいて二重分節解析器の有効性を示唆した。また、MDL を用いる従来法 [5] より高速であることが示された。sticky HDP-HMM は隠れ状態数を自動推定しつつ left-to-right の HMM を生成することも可能で動作生成までの繋げたモデリングも可能であると考えられる [6]。

3.2 自動車運転データ

Takenaka らは二重分節解析器を自動車運転データに適用し、二重分節点が単純な HMM による分節点よりも優位にドライバの意図変化点を予測しうることを示した [7]。この二重分節点に注目し動画の再生速度を変化させることで効果的な運転

動画要約が可能な事が示されている [8]。また、Taniguchi や Nagasaka らは NPYLM を未終端の符号列の教師なし形態素解析に拡張することで二重分節解析器に基づく予測器を提案した [9]。従来法より長い予測が可能になった。

4. まとめ

ノンパラメトリックベイズ二重分節解析器は未だ近似を含んでいるものの、実データから言語に至る教師なし学習モデルを示している。引き続き、言語獲得や動作獲得、意図変化点推定などを通じて記号創発ロボティクスへの貢献を目指したい。

参考文献

- [1] 谷口忠大. 二重分節構造に基づく教師なし学習と記号創発-時系列情報の分節化による記号創発ロボティクスへのアプローチ-. 人工知能学会誌, Vol. 27, No. 6, pp. 569-579, 2012.
- [2] D. Mochihashi, T. Yamada, and N. Ueda. Bayesian unsupervised word segmentation with nested Pitman-Yor language modeling. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1*, pp. 100-108, 2009.
- [3] T. Taniguchi and S. Nagasaka. Double articulation analyzer for unsegmented human motion using pitman-yor language model and infinite hidden markov model. In *Proceedings of 2011 IEEE/SICE International Symposium on System Integration*, 2011.
- [4] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky. The sticky hdp-hmm: Bayesian nonparametric hidden markov models with persistent states. Technical Report 2777, MIT Laboratory for Information and Decision Systems, 2007.
- [5] Tadahiro Taniguchi, Keita Hamahata, and Naoto Iwahashi. Unsupervised segmentation of human motion data using a sticky hierarchical dirichlet process-hidden markov model and minimal description length-based chunking method for imitation learning. *Advanced Robotics*, Vol. 25, pp. 2143-2172, 2011.
- [6] 古林邦彬, 谷口忠大. 手書き文字生成のための hdp-hmm による trajectory hmm のモデル選択. 第 13 回 計測自動制御学会 システムインテグレーション部門講演会, 2012.
- [7] Kazuhito Takenaka, Takashi Bando, Shogo Nagasaka, Tadahiro Taniguchi, and Kentaro Hitomi. Contextual scene segmentation of driving behavior based on double articulation analyzer. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012.
- [8] Kazuhito Takenaka, Takashi Bando, Shogo Nagasaka, and Tadahiro Taniguchi. Drive video summarization based on double articulation structure of driving behavior. In *ACM Multimedia*, 2012.
- [9] Tadahiro Taniguchi, Shogo Nagasaka, Kentarou Hitomi, Nainwala P. Chandrasiri, and Takashi Bando. Semiotic prediction of driving behavior using unsupervised double articulation analyzer. In *IEEE Intelligent Vehicles Symposium*, pp. 849-854, 2012.