

隠れマルコフモデルのための高速探索

背景・課題：隠れマルコフモデル（HMM; Hidden Markov Model）はシーケンスを確率に従った状態の遷移として表現するデータモデルです。音声認識、自然言語処理、遺伝子解析などの多くのアプリケーションにおいて使われています。

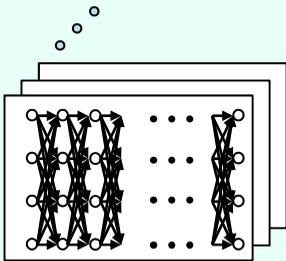
アプローチ：本研究では与えられた問い合わせシーケンスに対して、モデル集合の中から最適なモデルを探索する問題を対象とします。このような探索はモデルによって推定される尤度によって行われますが、多大な計算コストを必要とします。そこで尤度近似に基づく効率的な探索手法を提案します。

到達点：医療、遺伝子、交通量データを用いた実験により提案手法の有効性を確認しました。提案手法は探索漏れがないことを保証しながら、モデルの探索を高速に行うことができます。従来手法と比べ探索時間を大幅に低減化させ、最大で526倍の高速化を達成しました。

概要：

研究目的

問い合わせシーケンスについて、HMMの集合の中から尤度が高いモデルを探索。

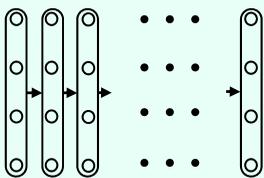


全てのモデルの尤度を求めるには高い計算コストがかかる。

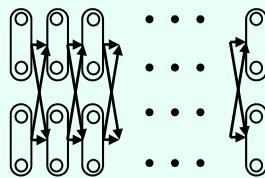
提案手法の特長

1. 高速探索
最適なモデルを効率的に探索。
2. 正確性
探索においては精度を犠牲にしない。
3. モデルタイプ
Ergodic HMM, Left-right HMMなどモデルタイプに依存しない。

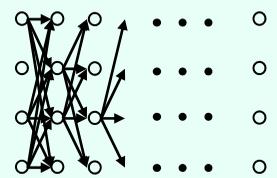
アプローチ



“類似した”モデルを統合することによりコンパクトなモデルを作成。



様々な粒度で近似尤度を計算。尤度の上限値を使ってモデルを枝刈り。



不必要なパスを取り除いて尤度計算をさらに高速化。

関連文献

Y. Fujiwara, Y. Sakurai, M. Yamamuro: “SPIRAL: Efficient and Exact Model Identification for Hidden Markov Models”, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 247–255, Las Vegas, Nevada, August 2008 (Best Paper Awards).
Y. Fujiwara, Y. Sakurai, M. Kitsuregawa: “Fast Likelihood Search for Hidden Markov Models”, ACM Transactions on Knowledge Discovery from Data (TKDD), Vol. 3, No. 18, November 2009.

連絡先

櫻井 保志 (Yasushi Sakurai)

協創情報部 知能創発環境研究グループ

