

## 二項事後分布の継承と W-B 平滑化に基づく 音声認識のための言語モデル

関西学院大学大学院 理工学部研究科  
情報科学科専攻 川端研究室 吉田 正太郎

音声認識における言語モデルの役割は、言語が発話された後に発話される発話の生成に制限を加えて探索空間を絞り込み、認識精度や認識速度を向上させることにある。現在の音声認識システムで多用されているのが n-gram 言語モデルである。n-gram 言語モデルは単語 n 個の共起連鎖に基づく統計的な言語モデルである。n-gram 言語モデルで重要となるのはゼロ頻度問題である。ゼロ頻度問題とは、学習データに中に出現しない単語連鎖(コンテキスト)の確率がゼロになってしまい、パープレキシティが無限大になる。この問題に対処するためにバックオフ平滑化という手法がよく用いられる。例えば、W-B(Witten-Bell)法、BPD(Binomial Posteriori Distribution)法などが挙げられる。W-B 法はコンテキスト毎に平滑化の程度を変える手法で、決め細やかな調整が可能である。BPD 法はベイズ法則の考えに基づき、学習データの量に対して適切な平滑化を行う手法である。W-B 法や BPD 法のパープレキシティの削減効果の評価により、両者には各々長所があり、うまく組み合わせることで相互的に働くことが予想された。W-B 法と BPD 法の確率計算の式には類似した式によって計算されている部分があり、本研究では、継承係数 の値を適切に設定することで両者の融合を試みた。これにより従来の BPD 法よりも決め細やかな確率推定が行える。

パープレキシティの評価では、結果は図のようになり。W-B 法よりも低いパープレキシティを得ており、改善がみられた。

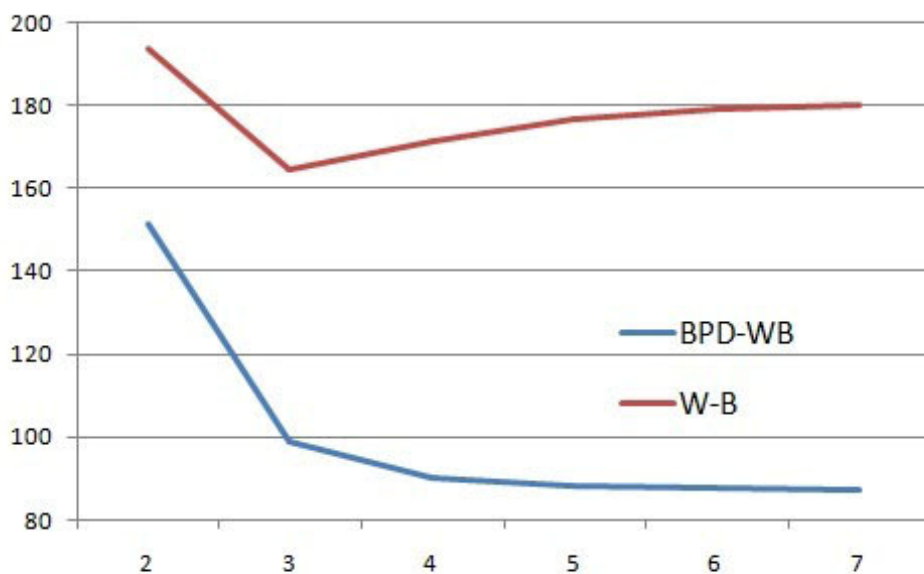


図 BPD-WB法とW-B法のパープレキシティによる評価