

音韻文脈の依存関係ネットワークに着目した 音韻 HMM の構成法

宮垣諒一[†] 川端豪[†]

潜在的ディレクレ配分法 (LDA) は大規模なテキスト検索のためのトピック処理の有効な手法である。LDA は大量のテキストデータを扱うことで潜在的なトピックのクラスタを探し出すことができる。本報告では LDA を音韻 HMM のクラスタリングに適用することを試みる。この枠組みでは、LDA におけるトピックをトライフォンのクラスタ、文書をトライフォン、単語をトライフォンの状態として対応付ける。クラスタの初期設定には既存のトップダウンおよびボトムアップクラスタリング手法を用いて、LDA で調整する。LDA 手法を音韻クラスタリングに用いることの有効性を、CSJ 講演データを用いて検討する。

Clustering of Phone HMMs based on Latent Dirichlet Allocation

Ryoichi Miyagaki[†] and Takeshi Kawabata[†]

The Latent Dirichlet Allocation (LDA) is a powerful method of topic handling for large scale text retrieval. The LDA finds out latent topic clusters by treating large amount of text data. This report try to adopt the LDA method to clustering of phone HMMs. In this framework, a topic corresponds to a triphone cluster. A document corresponds to a triphone. A word corresponds to a triphone state. Initial settings of triphone clusters are given using traditional top down and bottom up clustering techniques. And the LDA tunes them. Experimental results show that the LDA clustering is effective for CSJ lecture tasks.

1. はじめに

近年、計算機の高速化、大規模音声コーパスの整備などに伴い、大語彙連続音声認識に関する研究が盛んに行われている。大語彙連続音声認識においては、音韻を単位として HMM (Hidden Markov Model) を学習することが一般的である。さらに音韻の音響の特徴は先行及び後続音韻の影響を大きく受けるため、3 音韻の連鎖 (トライフォン) を単位とする HMM を学習することがよく行われている。

しかし、トライフォンには膨大な種類が考えられるため、学習データ中に十分な個数のトライフォンが含まれない状況が起こりうる。学習データに現れないトライフォンは原理的に認識できず、また出現回数が少なければ HMM の信頼性が低下するという問題が起きる。

この問題を解決するために、複数のトライフォンを集めてクラスタを作り、互いに代用することで見かけ上の学習データを増やす手法が研究されてきた。音韻文脈に基づくトップダウン的な情報を用いてクラスタリングを行う TBC (Tree Based Clustering) が有名である [1][2][3]。TBC では HMM を構成する状態ごとにクラスタリングを行う。ある状態に対し、同じ中心音韻を持つトライフォン全体を大きな一つのクラスタと考え、音韻文脈や学習される HMM の統計的信頼性を手がかりにクラスタを分割していく。一つのクラスタに含まれるトライフォンはその状態を共有する。この操作を HMM の各状態について行う。その他にも、音響的に類似するトライフォンを探しまとめ上げるボトムアップクラスタリング [4] や、トップダウンとボトムアップを併用するクラスタリング手法 [5] が報告されている。

これらの音韻クラスタリングの構造に着目すると、トライフォンの各状態やクラスタをノード、データの共有関係をリンクとするネットワーク構造であることが分かる。本研究ではネットワーク構造に基づいて各トライフォンの状態を自動的にクラスタに割り当てる手法として、文書クラスタリングでよく用いられる LDA (潜在的ディレクレ配分法, Latent Dirichlet allocation) [6] の利用を試みる。LDA による音韻クラスタリングを行うことで、より高精度な音響モデルを学習する。

2. 音韻 HMM クラスタリング

2.1 クラスタリングの必要性

大語彙連続音声認識では音韻 (phoneme) 単位で HMM を作る事が一般的である。例えば、連続音声中の「秋」という単語を認識するためには、/a/, /k/, /i/ という各音韻の HMM を連結して確率を計算すればよい。しかし、音韻の音響的特徴は前後の音

[†] 関西学院大学 理工学研究科
School of Science and Technology, Kwansai Gakuin University

韻によって大きく影響を受ける。同じ音韻/k/でも「秋 (/a k i)」の/k/と「駅 (/e k i)」の/k/では先行音韻が異なるので舌及び顎の位置が異なり、音響的に違いが現われる。そこで、音韻環境を考慮した認識単位として、3音韻連鎖(トライフォン)が用いられる。

認識単位としてトライフォンを用いる際の問題点は、その種類の多さである。例えば日本語の場合、音韻の種類数を仮に40とすると、トライフォンの総数は64000(40×40×40)もの数になってしまう。このためトライフォンあたりの学習データが少なくなりHMMの統計的信頼性が低下する。最悪の場合、学習データに現れなかったトライフォンについては学習が行えない。

そこで、いくつかのトライフォンをグループ化(クラスタリング)し、一つのHMMを学習することによってモデル数を削減することが考えられる。例えば、トライフォンを音韻文脈の類似性に従って段階的に分割するトップダウン的考え方に基づくグループ化手法が研究されている。音響的に類似するトライフォンをボトムアップ的にまとめていく考え方[4]や、またトップダウンとボトムアップ両方のクラスタリング法を組み合わせた手法[5]が挙げられる。

本研究では、与えられた分量の音声データを用いて大語彙連続音声認識のための最良の音響HMMを学習するため、トライフォン・クラスタをノードとし、データ共有関係をリンクとするネットワークに基づくHMMパラメータの共有構造を提案する。また今回各トライフォン・状態のクラスタへ割り当てる方法として、文書クラスタリング手法の一種であるLDAを用いて、最適なクラスタ・重みを自動決定してくれるようなネットワークモデルを構築することで、大語彙連続音声認識のための最良のHMM学習を目指す。

2.2 Tree Based Clustering

Tree Based Clustering(TBC)は、先行・後続音韻を考慮し音響的に類似した環境を木構造で構成してトップダウン的にクラスタ化を行う手法である[1][2][3]。TBCの分割条件は、音声の特徴に基づく先行音韻・後続音韻に関する2者択一の質問である。たとえば「先行音韻が破裂音であるか?」などというような質問があり、すべてのトライフォンはYes/Noのいずれかに分類される。手順として、まず中心音韻が共通の全てのトライフォンを一つの集合とし、これをルートノードとする。次に、分割条件の一つ一つに従ってノードを仮分割する。それぞれの仮分割に対して学習されるHMMの尤度などの評価尺度の分割前後の変化量を求め、これが最大となる分割条件を選択してノードを分割する。この分割を繰り返すことによって決定木を生成する。最終的なリーフノードがクラスタになる。

図1に、中心音韻/a/に対し、先行音韻あるいは後続音韻の音韻論的類似性に基づいて場合分けを進める手続きを示す。クラスタリング後、1つのリーフノード(最末

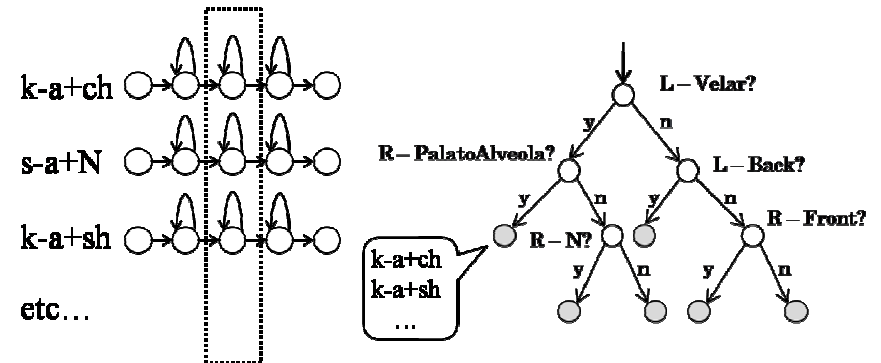


図1 Tree Based Clustering

端のノード)に含まれるトライフォンは1つのHMMの状態を共有する。以上の操作により、出現頻度の低いトライフォンに対しても十分な学習データを確保できる。また学習データに出現しないトライフォンも決定木を辿ることでいずれかのクラスタに含まれるため、HMMを代用することができる。

2.3 LDA クラスタリング

2.3.1 LDA の概要

LDA[6]はBleiらによって提案された、多重トピック文書の生成過程を階層ベイズモデルによりモデル化した確率モデルである。LDAモデルは、文書からトピックが生成される事象をディリクレ分布に従う確率として与える。

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^K \alpha_i)}{\prod_{i=1}^K \Gamma(\alpha_i)} \prod_{i=1}^K \theta_i^{\alpha_i - 1} \quad (1)$$

$$\alpha_i > 0$$

$$\theta_i > 0, \sum_{i=1}^K \theta_i = 1$$

ここで、 α はディリクレ分布のパラメータであり、 K 次元ベクトルである。また θ は各トピックにおける多項分布のパラメータであり、文書が各トピックを生成する確率を示す。単語のユニグラム確率を $\beta = (\beta_{11}, \beta_{12}, \dots, \beta_{1V}, \beta_{21}, \dots, \beta_{KV})$ と表す。 β_{ij} はトピック i で単語 j が出現する確率であり、 V は単語数である。

LDAは以下のように多重トピック文書の生成過程をモデル化している。

1. Choose $\theta \sim Dir(\theta|\alpha)$
2. For each of the N words W_n
 - a) Choose a topic $Z_n \sim Multinomial(\theta)$
 - b) Choose a word W_n from $p(W_n|Z_n, \beta)$

まず、ディリクレ分布 $Dir(\theta|\alpha)$ から各トピックの出現確率を示す θ がサンプリングされる。次に、各文書から生成する単語の数 ($n=1, 2, \dots, N$) だけ以下の操作を繰り返す。まずは多項分布 $Multinomial(\theta)$ から n 番目に生成される単語のトピック Z_n がサンプリングされる。そしてトピック Z_n において単語 W_n が生成される確率 $p(W_n|Z_n, \beta)$ に従い W_n がサンプリングされる。これで、単語とそのトピック対 (W_n, Z_n) が生成される。

LDA は、パラメータを α, β として以下のように定式化される。

$$p(W|\alpha, \beta) = \int \frac{\Gamma(\sum_{i=1}^K \alpha_i)}{\prod_{i=1}^K \Gamma(\alpha_i)} \prod_{i=1}^K \theta_i^{\alpha_i - 1} \left(\prod_{n=1}^N \sum_{i=1}^K \beta_{iW_n} \theta_i \right) d\theta \quad (2)$$

2.3.2 変分 EM 法による LDA モデルのパラメータ推定

LDA では変分 EM 法を用いて、式(2)における事前分布のパラメータ α と β を推定するための更新式を導出する。 α の導出には、定点反復法[7]を用いた。定点反復法では式(3)に従いパラメータを更新させる。また式(3)はディガンマ関数の逆算数を計算することで α を導出する。即ち、 $\Psi^{-1}(\Psi(\alpha_z^{new}))$ を求めることになる。式(4)はディガンマ関数の逆関数であり、式(5)は β の更新式である。 $\delta(x, y)$ はクロネッカーのデルタ関数で、 $x = y$ の時 1, $x \neq y$ の時 0 を返す。

(M-Step)

$$\Psi(\alpha_z^{(t+1)}) = \Psi\left(\sum_z \alpha_z^{(t)}\right) + \frac{\sum_d (\Psi(\gamma_{dz}) - \Psi(\sum_z \gamma_{dz}))}{D} \quad (3)$$

$$\Psi^{-1}(y) \approx \begin{cases} \exp(y) + 1/2 & \text{if } y \geq -2.22 \\ -1/(y - \Psi(1)) & \text{if } y < -2.22 \end{cases} \quad (4)$$

$$\beta_{zw} \propto \sum_{d=1}^D \sum_{n=1}^{N_d} \varphi_{dnz} \delta(d_n, w) \quad (5)$$

次に、 α と β の更新式に含まれる 2 つの変分パラメータ γ と φ を導出する。 γ_d は文書 d でのディリクレ分布のパラメータで K 次元ベクトルである。また φ_{dni} は、文書 d で n 番目の単語のトピックが i である確率である。学習時には α と β を推定する前に、 γ と φ を

交互に収束するまで更新することで、変分事後分布を最大化し、真の事後分布に近づけていく必要がある。 γ と φ は式(6)、式(7)によって更新される。

(E-Step)

$$\gamma_{di}^{(t+1)} = \alpha_i^{(t)} + \sum_{n=1}^N \varphi_{dni}^{(t+1)} \quad (6)$$

$$\varphi_{dni}^{(t+1)} \propto \beta_{iw}^{(t)} \exp\left(\Psi(\gamma_{di}^{(t+1)}) - \Psi\left(\sum_{j=1}^K \gamma_{dj}^{(t+1)}\right)\right) \quad (7)$$

2.3.3 LDA に基づく音韻クラスタリング

LDA を文書クラスタリングとして用いる際には、文書を単語の集合と考えて、各単語を生成したトピックを推定する。本研究では LDA を音韻クラスタリングに適用することから、LDA における文書をトライフォン、単語をトライフォンの各状態、トピックをクラスタとして扱うことでクラスタリングを行う。これより、HMM の各状態は第 2 状態、第 3 状態、第 4 状態に関わらず同じクラスタを共有する。クラスタリングは TBC と同様、同じ中心音素を持つトライフォンごとに行う。また今回は出現頻度が高く、認識精度に大きく影響する 5 母音のみのクラスタリングを行う。

初期パラメータ α に関して、あるクラスタ α_i が学習データ 1 講演中に出現する回数を設定する。

LDA の単語ユニグラム確率ベクトル β は、今回の音響モデルに対する適応において、トライフォンの各状態がどのクラスタに関係付けられる確率に対応する。初期値として、0.5 から離れた値 (変位) を与えることにより、クラスタの初期バリエーションを設定する。初期バリエーションに関して、我々が過去に検討したトップダウン及びボトムアップ手法によるクラスタリングの結果を利用する。あるトライフォンが既存手法のあるクラスタに含まれる時、確率ベクトル β の対応する要素に変位を与える。例えば変位を 0.8 とすると、クラスタに含まれない時の値は 0.2 となる。これを、(0.8 : 0.2) と記す。

まず TBC のクラスタに基づいて β の初期値を設定する検討を行う。このときの変位は、1.0, 0.8 に加え、比較のために 0.5 (変位なし) の 3 通りについて検討する。

続いて、トライフォンの状態パラメータ間の分散正規化距離に基づくボトムアップクラスタ及び TBC に基づくトップダウンクラスタを用いて β の初期化を行う。この時の変位は、例えばトライフォンがボトムアップクラスタに含まれる時 0.5, TBC クラスタに含まれる時 0.4, それ以外の時 0.1 といった与え方をする。これを (0.5 : 0.4 : 0.1) と記す。

表1 実験条件

基本周波数	16kHz
分析窓	Hamming 窓
分析窓長	25ms
フレーム周期	10ms

LDA における文書中の単語数は、音韻クラスタリングではトライフォンの状態数なので 3、学習用文書数は学習データ中に出現するトライフォンの種類数とする。トピック数（クラスタ数、ディリクレ分布の次元数）は、TBC でクラスタリングされた各状態数の合計である。

LDA によるクラスタ作成方法として、LDA によって取り出された組（状態・クラスタ）をもとに、クラスタ内で一定の共起回数より多い状態のみを残すことで一つのクラスタを生成する。ある一つの状態が複数のクラスタに出現する場合は、共起回数の多いクラスタを選択する。同じ共起回数の場合はランダムにクラスタを選択する。

3. 評価実験

3.1 実験条件

学習データとしては、日本語話し言葉コーパス (CSJ) [8] の中から、学会講演 100 セットを選択し用いる。ただし、話者はすべて女性とした。各講演は 10 分から 20 分程度である。特徴量は 12 次元 MFCC と対数エネルギー、及びそれらのデルタ項、デルタデルタ項の計 39 次元である。トライフォンの混合数は 16 とする。学習ツールとして、HTK[9] を用いる。

評価データには学習データに含まれない学会講演 10 講演分の女性話者データを使用する。音声認識デコーダは Julius[10] を用いる。その他実験条件を表 1 に示す。認識結果は正解文章に対する単語単位での単語正解精度 (Acc)、単語誤り率 (WER) について (8) 式・(9) 式に定義し、集計した。ここで正解文の単語数を N 、認識結果における置換誤り単語数を S 、挿入誤り単語数を I 、脱落誤り単語数を D とする。

$$ACC = \frac{N - S - D - I}{N} * 100.0 \quad (8)$$

$$WER = \frac{S + D + I}{N} * 100.0 \quad (9)$$

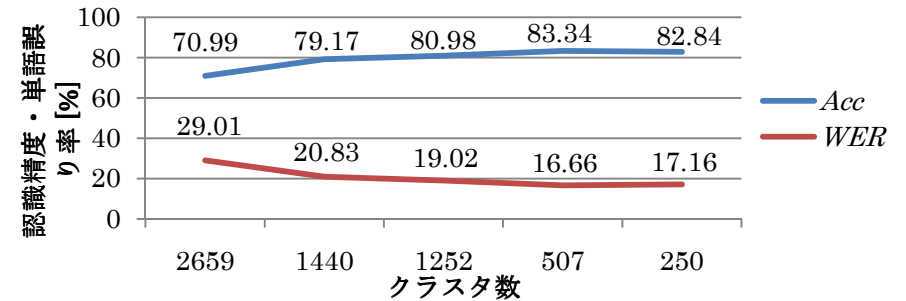


図2 トップダウン的に生成したクラスタ(TBC)に基づく音響モデルを用いた音声認識実験の結果

3.2 トップダウン的に生成したクラスタ (TBC) に基づく音響モデルを用いた音声認識実験

まず、トップダウン的手法によって生成したクラスタに基づく音響モデルの性能を把握するために、TBC で様々な数のクラスタを生成し、音響モデルを学習して音声認識の性能を調べた。実験結果を図 2 に示す。

実験結果より、Acc・WER 共にクラスタ数が 507 のところで精度がピークに達し、それ以降スコアは低下する。これより TBC によって状態数が約 500 までは状態の共有化が上手く行われているが、その後は音響的に類似しない状態まで共有化がされ認識精度に影響されているのではないかと考えられる。

3.3 LDA を利用して作成したクラスタを加えた音響モデルでの音声認識実験

3.3.1 LDA 実験条件

TBC で作成したクラスタに、LDA によって作成したクラスタを加えた音響モデルを用いて実験を行う。初期パラメータとして使用される TBC クラスタは、認識結果が一番良かったクラスタ数 507 のものを使用する。

初期パラメータ β に関して、2.3.3 節に従って 2 通りの実験を行う。まず一つ目の実験は、あるトライフォンが TBC のあるクラスタに含まれる時の初期パラメータ β の対応する要素に変位を (1.0 : 0.0), (0.8 : 0.2), (0.5 : 0.5) の 3 種類で設定した。

二つ目の実験は、トップダウンとボトムアップクラスタリングを併用した方式を初期パラメータ β に適用する。その時の変位を、(0.5 : 0.4 : 0.1), (0.4 : 0.3 : 0.3) の 2 種類で設定した。

表2 LDAによって取り出されたクラスタ(トピック)と
 トライフォンの状態の共起回数 (上位8個)

a_s24_LDA		a_s26_LDA		a_s28_LDA	
トライフォンの状態	共起回数	トライフォンの状態	共起回数	トライフォンの状態	共起回数
r-a+m2	9	ts-a+sh2	9	k-a+a2	10
r-a+a2	8	t-a+o4	8	k-a+tj2	9
u-a+p2	8	s-a+a2	7	k-a+p2	8
d-a+d4	7	s-a2	7	g-a+ts3	8
i-a+k3	7	z-a+ch3	7	k-a+q2	8
r-a+i2	6	z-a+sh2	7	y-a+s2	8
r-a+e2	6	t-a2	6	d-a+sh3	8
r-a+ch2	6	ts-a+e2	6	i-a+n4	8

変分 EM 法の演算回数は、E-Step を 10 回、M-Step を 15 回、全体の収束判定は、パラメータ α の合計値の変化量が 1.0 を切ったところでストップするように設定した。また LDA によって取り出される組 (状態・クラスタ) の数は、学習データのトライフォン数によって決まるが、取り出される組数に差をつけるようトライフォン数 $\times 5$ 、 $\times 10$ と倍増することでより多くの組数を取り出し、共起回数に差をつけるようにする。クラスタに一定回数以上出現した状態のみをグループ化するが、倍増する学習トライフォン数によってその閾値を変えた。

3.3.2 LDA によるクラスタリング

LDA によってクラスタリングされた結果例を表 2 に示す。表 2 は中心音素 /a/ の場合であり、ここではトライフォンの状態がトピック (LDA クラスタ) に共起した回数上位 8 個を挙げた。 β の初期値設定は TBC の知見のみを反映させた。「a_s24_LDA」はトピックを、状態欄の「r-a+m2」はトライフォン r-a+m の第 2 状態を示している。表 2 のように、LDA によるクラスタリングでは同じトピック内に第 2 状態、第 3 状態、第 4 状態関係なく状態選出が行われている。また、先行・後続音韻が類似するトライフォンが同じクラスタに多く共起することから、 β の初期値である TBC の情報が LDA に影響されていることが分かる。

ここで、トピックとの共起が 4 回以上、5 回以上などのように一定回数以上であるトライフォンの状態を LDA によるクラスタリングにおける新たなクラスタとする。

表3 LDA クラスタリングの β の初期化に TBC を利用した場合の音声認識実験結果

β の初期値に与える変位 (TBC:なし)	LDA クラスタ数	Acc[%]	Err[%]
TBC (LDA なし)		83.34	16.66
LDA (1.0:0.0) (TBC で初期化)	207	82.73	17.27
LDA (0.8:0.2)	67	83.75	16.25
LDA (0.5:0.5) (初期化なし)	32	83.04	16.96

表4 LDA クラスタリングの β の初期化にボトムアップクラスタリングと TBC を併用した場合の音声認識実験結果

β の初期値に与える変位 (Bottom:TBC:なし)	LDA クラスタ数	Acc[%]	Err[%]
TBC (LDA なし)		83.34	16.66
LDA (0.5:0.4:0.1)	82	83.41	16.59
LDA (0.4:0.3:0.3)	12	83.5	16.5

3.3.3 LDA クラスタリングの β の初期化に TBC を利用した場合の音声認識実験

β の初期化に TBC を用いた場合の LDA クラスタリングによる音響モデルでの音声認識実験を行った。学習トライフォン数や、クラスタ化する共起回数などのパラメータ設定をいくつか変化させて実験を行ったが、比較のため認識精度が一番向上したものを挙げる。表 3 に実験結果を示す。 β の初期値の変位が (0.8:0.2) の時は、Acc=83.75%、Err=16.25% の性能が得られたが、これは TBC のみの実験で精度が一番良かったクラスタ数 507 の Acc=83.34、Err=16.66 よりも向上した。また β の初期値において (1.0:0.0) (TBC で初期化) や (0.5:0.5) (初期化なし) の条件よりも、中間的な変位を与えるほうが結果は良くなる。実験では (0.8:0.2) が最良の性能を示した。

3.3.4 LDA クラスタリングの β の初期化にボトムアップクラスタリングと TBC を併用した場合の音声認識実験

β の初期化にボトムアップクラスタリングと TBC を併用した場合の LDA クラスタリングでの音声認識実験を行った。実験結果を表 4 に示す。 β の初期値の変位が (0.5:0.4:0.1)、(0.4:0.3:0.3) 両方共に TBC のみ (クラスタ数 507) の結果よりも性能が向上した。表 3 の β 初期値変位 (0.8:0.2) と表 4 (0.5:0.4:0.1) を比較するとボト

ムアップクラスタリングの要素を導入したことによって認識精度に向上は見られなかったが、表 3 (0.5 : 0.5) と表 4 (0.4 : 0.3 : 0.3) を比べると Acc, Err 共にスコアは上がった。ボトムアップクラスタのみを β の初期化に用いたことで精度が向上したと考えられる。

4. おわりに

文書クラスタリングに用いられる LDA を音響クラスタリングに利用することを考えた。LDA の単語に対応する確率ベクトル β を、トピックとトライフォンの状態に対応させ、その初期化の検討を行った。またこの手法では変分 EM 法を用いて、パラメータ α と β を推定するための更新式の導出を行った。

LDA で作成したクラスタを加えた音響モデルを用いた音声認識実験では、 β の初期値に与える変位を (0.8 : 0.2) と設定した場合、TBC で一番性能が良いものより精度は向上した。初期値 β にボトムアップクラスタリングの知見を入れた場合の認識実験では、 β の初期値に与える変位を (0.5 : 0.4 : 0.1), (0.4 : 0.3 : 0.3) と設定したが、両方とも TBC のみ (クラスタ数 507) の結果よりもスコアは向上した。また初期パラメータ β に最適な設定方法があると考えられる。以上のことから LDA による音響クラスタリングの有効性が確認できた。

参考文献

- 1) S.J.Young, J.J.Odell, P.C.Woodland : Tree-Based State Tying for High Accuracy Acoustic Modelling , ARPA Workshop on Human Language Technology, pp.307-312 (1994)
- 2) 嵯峨山 茂樹 : 音素環境クラスタリングの原理とアルゴリズム, 電子情報通信学会技術報告, SP87-86, pp. 1-8 (1987)
- 3) Sung-Il Kim, Tetsuro Kitazoe : Continuous Speech Recognition Using Tree based State Tying , 情報学研報, 98-SLP-20-16 (1998)
- 4) X. Aubert, P. Beyerlein, M. Ullrich: A Bottom-Up Approach for Handling Unseen Triphones in Large Vocabulary Continuous Speech Recognition, Proceedings ICSLP, Philadelphia, PA, U.S.A., pp. 14-17 (1996)
- 5) 宮垣諒一, 川端豪: トップダウン及びボトムアップ手法に基づく音韻 HMM のクラスタリング, Vol.2010-SLP-82 No.12
- 6) David M.Blei, Andrew Y.Ng, Michael I.Jordan : Latent Dirichlet Allocation, Journal of Machine Learning Research, vol.3, 993-1022(2003)
- 7) T.Minka : Estimating a Dirichlet distribution, <http://www.stat.cmu.edu/minka/papers/dirichlet.html>
- 8) 前川喜久雄 : 「日本語話し言葉コーパス」 付属ドキュメント 「日本語話し言葉コーパス」の概観(2004)
- 9) S. Young, et. al. : The HTKBook , Entropic Cambridge Research Laboratory
- 10) 河原達也, 李晃伸 : 連続音声認識ソフトウェア Julius, 人工知能学会誌, Vol.20, No.1, pp41-49(2005)
- 11) 岡登洋平, 速水悟, 板橋秀一 : クラスタリングによる HMM 間の距離尺度の検討, 信学技報, SP94-16, pp.15-20 (1994)
- 12) 正田備也, 喜安千弥, 宮原末治 : 文書クラスタリングのための潜在的ディリクレ配分法による次元圧縮, 情報処理学会研究報告, 2007-DBS-143(64), pp.381-386
- 13) Jose B.Marino, Albino Nogueiras, Antonio Bonafonte : The demiphone: an efficient subword unit for continous speech recognition, Proc.EUROSPPECH97, pp.1215-1218