

# 階層ディリクレ過程隠れマルコフモデルを用いた 非分節動作模倣学習手法

谷口忠大 濱畑慶太\* (立命館大学) 岩橋直人 (情報通信研究機構)

## Imitation Learning Method for unsegmented motion using Hierarchical Dirichlet Process Hidden Markov Model

\*T. Taniguchi, K. Hamahata (Ritsumeikan University) and N. Iwahashi (NICT)

**Abstract**— Natural unsegmented human motion data contains various types of elemental motions. When an observer imitates a type of motion from such data, the learner has to segment and extract a unit motion from the data. To achieve this, ergodic hidden Markov model is often used to model the unsegmented human motion. However, to set a proper number of hidden states is difficult in order to model unseen various types of motions. In addition to that, EM algorithm which is usually used to estimate ergodic hidden Markov model's parameters is tend to easily caught to local minima. To overcome these problems, we utilize sticky hierarchical Dirichlet process hidden Markov model which is a type of a nonparametric Bayes model. An experiment is performed to evaluate its capability.

**Key Words:** unsegmented human motion, imitation learning, nonparametric Bayes model, sticky hierarchical Dirichlet process hidden Markov model (sHDP-HMM), ergodic HMM

### 1 はじめに

人間の模倣学習能力の高さが、近縁種からの違いを特徴付けると言われる<sup>1)</sup>。模倣能力は身体的技能を身につけるためのみならず、ジェスチャをはじめ社会的なサインやルールを身につける為にも必要であり、記号接地や記号創発の視点からも重要な研究課題である<sup>2)</sup>。知能ロボット研究においては模倣学習を教師有り学習として矮小化する議論も見受けられるが、模倣学習は本質的により多くの知的活動を含んでおり、単純に教師有り学習として議論が完結するものではない。谷口らは役割交代模倣を数理的にモデル化し、実時間での単純なインタラクションから役割交代を含め、模倣学習を進めるプロセスが機械学習の定式化としてはEMアルゴリズムに基づく教師なし学習の形に定式化されることを指摘している<sup>3, 4)</sup>。また、一歳程度の幼児であっても、他者の連続的な動作から動作単位を切り出して模倣することが出来る。これを本稿では非分節動作からの模倣学習と呼び、研究の対象とする<sup>5)</sup>。

非分節動作からの模倣学習のためには、人間の視点から見て、もしくはロボットの何らかの実用主義的な視点から見て、単位となり得る動作を抽出するという分節化処理と、連続的にあたえられ、様々な動作要素を含んだ非分節な動作時系列をモデリングするという二つの学習プロセスを同時に行う必要がある。非分節動作からの模倣学習には複数のアプローチが存在するが、主なものにRNNのような非線形予測機を用いて単位動作表現を行うものと、線形モデルやガウス分布をより細かな動作単位に区分しこれらの遷移列により単位動作を表現するものがある。前者ではRNNPBやmixture of RNN experts, RNN-SOM, DBSOMなどが提案されている<sup>6, 7, 8)</sup>。後者ではHMMやSwitching AR Model(SARM)などを用いるアプローチがあ

る<sup>15, 9)</sup>。本稿では、後者のアプローチに着目する<sup>2)</sup>。後者の手法の概略としては、まず離散個の隠れ状態とそれに対応する確率分布<sup>3)</sup>により動作時系列をモデリングし、隠れ状態を表すインデックス列に符号化する。この符号列から、繰り返し現れる単語に相当する文字列を抽出することで、単位動作の抽出を行う。

HMMやSARMを模倣学習に用いる場合、分節化された動作に対しては始点と終点が明確であり、また、何と何が同じ動作時系列であるかということが、しばしば自明であるので、HMMの構造としては有限個の隠れ状態を一方向的に遷移するleft-to-right HMMを用いる事が有効であり、実際に用いられている<sup>10)</sup>。しかし、一方で非分節動作の模倣学習では単位動作の始点と終点は明確でない。そのために、通常は全ての隠れ状態に遷移しあうことを前提としたergodicなモデルを用いる事になる。しかし、ergodicなモデルでは、モデルの自由度が高いために、EMアルゴリズムのような局所探索法では、初期値に依存し豊富な局所解に捕われてしまう問題がある。また、分節化された動作の学習に比べると、何種類の単位動作がどのように時系列中に含まれているのかを、事前に把握することがより困難なために、隠れ状態数を事前に設計することがより困難である。

隠れ状態数の決定のためには、MDLを始めとしたモデル選択手法を用いる事が考えられるが、多くのモデル選択手法ではオフラインでの計算が必要になると同時に、隠れ状態に応じた再学習が必要となる。これにより、計算量が膨大となると同時に、ロボットによるオンラインでの適応学習がより困難になってしまう。そこで、隠れ状態をオンラインで自動決定するとともに、局所最適値に陥ることのないモデル化手法が必要

<sup>1)</sup>MOSAICによるモデルもしばしば引用されるが、線形モデルでの分割のみでは、通常ここでいう単位動作にはならない。よって比較対象からは除外する。

<sup>2)</sup>現時点ではどちらのアプローチも一長一短の面があり、両者のアプローチが研究されるべきである。

<sup>3)</sup>主に線形ダイナミクスもしくは状態空間上のガウス分布

\*現在、アクセンチュア(株)に勤務

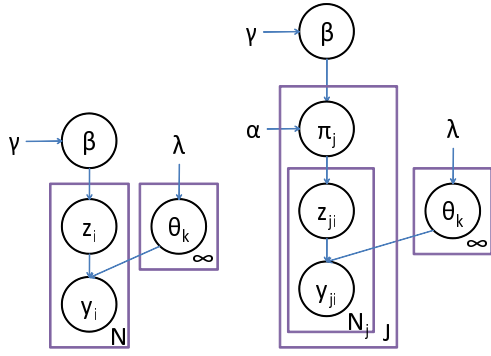


Fig. 1: Graphical model of Dirichlet Process Mixture

となる．本稿では近年注目されているノンパラメトリックベイズ法<sup>11)</sup>に着目し，その一種である sticky HDP-HMM<sup>12)</sup>を非分節動作模倣のための動作モデリングに利用し，その有用性を検証する．

## 2 ノンパラメトリックベイズ法

ノンパラメトリックベイズ法は階層型ベイズ推論でありながら，問題の複雑性に依りてモデルの複雑性を自動決定することの出来る機械学習手法である．本章では，ディリクレ過程をはじめに，本稿で用いる sticky HDP-HMM までを概説し，sticky HDP-HMM のアルゴリズムを示す．

### 2.1 ディリクレ過程

音声や動作のモデリング，または，様々な状態量のクラスタリングに混合分布が用いられる．しかし，尤度最大化基準に従い EM アルゴリズムを用いるにせよ，ベイズ的な手法をとるにせよ，隠れ状態数をいくつにするかというモデル選択の問題がつきまとう．しばしば MDL が代表格であるモデル選択手法を用いて複雑さと尤度のバランスから適切な隠れ状態数が選択される．また，クロスバリデーションにより汎化誤差の視点から隠れ状態数が選択される．しかし，これらの手法はオフラインであると同時に計算量も多くなる．これに対し，ディリクレ過程では隠れ状態自体の生成過程をモデル化することにより，必要な数だけの隠れ状態が自然と決定されるという仕組みを持つ．ディリクレ過程<sup>13)</sup>は無数の確率分布を生成する過程であり，無限次元のディリクレ分布を構成していると考えられる．ディリクレ過程のグラフィカルモデルを Fig.1(left) に示す．

通常のパラメトリックベイズモデルの混合分布では，有限次元のディリクレ分布から多項分布が生成され，その多項分布に従い隠れ状態が選択され，その隠れ状態に相当する出力分布<sup>4)</sup>から観測変数が出力されるという生成モデルが仮定される．これに対し，ディリクレ過程を事前分布としたディリクレ混合過程 (Dirichlet Process Mixture: DPM) では，観測データに応じて隠れ状態の数，つまりクラスターの数決定される<sup>5)</sup>．

Sethuraman<sup>14)</sup>は，ディリクレ過程の構成法として Stick Breaking Construction を提案した．Stick Breaking Construction は長さ 1 の棒をベータ分布から出力

<sup>4)</sup>混合ガウス分布の場合はガウス分布

<sup>5)</sup>実際にはディリクレ過程は無数次元のディリクレ分布であり，その中で有限個の隠れ状態のみが結果的に用いられるという表現の方が妥当であると，筆者は考える．

される値により，次々と分割することで無限次元の多項分布を構成する．

$$v_k | \gamma = \text{Beta}(1, \gamma) \quad k = 1, 2, \dots \quad (1)$$

$$\beta_k = v_k \prod_{l=1}^{k-1} (1 - v_l) \quad k = 1, 2, \dots \quad (2)$$

上式により，無限次元の多項分布  $\beta$  が形成される．これは  $\beta \sim GEM(\gamma)$  と記述される．この他にディリクレ過程の表現としては中華飯店過程 (Chinese Restaurant Process: CRP) がある．CRP は店に入ってきた  $i$  番目の客が，すでに座っている人の数に比例した確率でテーブルに座り，また， $\frac{\gamma}{i-1+\gamma}$  の確率で新しいテーブルに座るというルールによってディリクレ過程を表現する．CRP の重要な性質として交換可能性 (exchangeability) があり，これによってギブスサンプリングとの親和性が非常に高くなっている<sup>11)</sup>．

DPM では選択された隠れ状態に対応した出力分布パラメータ  $\theta_k$  がその事前分布から生成され，その出力分布から観測値  $y_i$  が出力される．

### 2.2 階層ディリクレ過程

階層ディリクレ過程 (Hierarchical Dirichlet Process) では，ディリクレ過程により生成された確率分布がさらに，基底分布となることで階層化される．階層ディリクレ混合過程 (Hierarchical Dirichlet Process Mixture: HDP) のグラフィカルモデルを Fig.1(right) に示す．CRP に対し，HDP では中華飯店フランチャイズ (Chinese Restaurant Franchise: CRF) により生成過程が表現される．CRF では  $J$  個のレストランがあり，全てのレストランが一つの HDP グループに所属している．全てのレストランは無数に多くのテーブル (クラスターに相当) がある．それらのテーブルに客 (観測変数に相当) が座る． $j$  レストランの  $i$  番目の客は  $t_{ji} \sim \pi_j$  に座り，その机には  $k_{jt} \sim \beta$  の皿 (出力分布のパラメータに相当) が割り当てられる．観測変数  $y_{ji}$  はグローバルな変数  $\theta_{ji}^* = \theta_{jt_{ji}}^* = \theta_{k_{jt_{ji}}}$  から生成される．HDP は，それぞれアトム (atom) を共有した  $J$  個のディリクレ過程を生成する．

### 2.3 階層ディリクレ過程隠れマルコフモデル

上記，階層ディリクレ過程のアイデアを隠れマルコフモデルに拡張したものが Teh ら<sup>15)</sup>によって提案されている階層ディリクレ過程隠れマルコフモデル (Hierarchical Dirichlet Process - Hidden Markov Model: HDP-HMM) である．HMM に無限次元の潜在状態を仮定することにより，データに応じて潜在状態の数を決定するような柔軟なモデルになる．HDP-HMM のグラフィカルモデルを Fig. 1(top) に示す<sup>6)</sup>．

集中度パラメータ  $\gamma$  をパラメータとする  $GEM(\gamma)$  によって  $\beta$  が生成され，これに基づいて  $\pi_k$  が生成される．これらは隠れ状態に対応した多項分布であり，現在の状態に依存して次の状態を多項分布で選択するために，無限次元の状態遷移行列の意味を持つ．また，これらの隠れ状態に応じて，出力分布のパラメータ割り当てられる．HDP-HMM では，事前に left-to-right

<sup>6)</sup>HDP-HMM は無限の隠れ状態を持つ HMM ということで infinite HMM と呼ばれる場合もある<sup>16)</sup>．

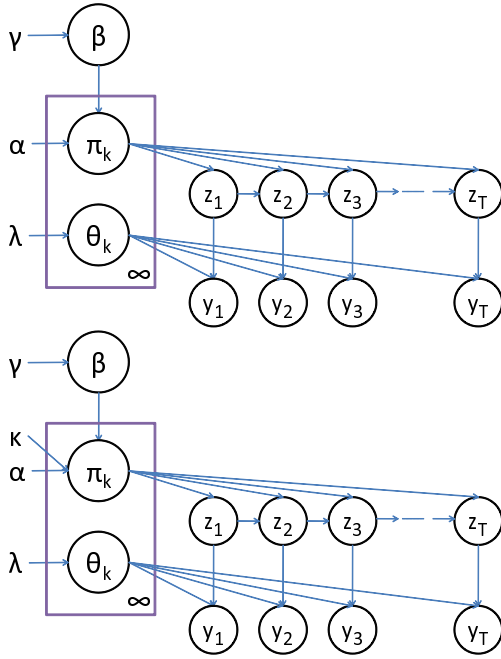


Fig. 2: Graphical model of (top) HDP-HMM and (bottom) sticky HDP-HMM

などの構造を仮定しないために、ある潜在状態に対して全ての潜在状態から遷移する可能性になり、基本的に HDP-HMM は ergodic HMM となる。しかしながら、HDP-HMM は生成される  $\pi_k$  に制約が無いことから、隠れ状態間を頻繁に切り替わる事になる。これに対して、音声や動作のモデリング、分節化といった HMM の応用上では、時間的にある程度十分な長さ一つの状態を維持するかが望ましい事がしばしばある。従来の HMM では自己遷移確率の初期値を高める、出力分布の初期値配置を工夫するなどの手段がとられる場合が多い。これに対して、Fox らは<sup>12)</sup>HDP-HMM に対して自己遷移確率を高めるハイパーパラメータを付加した sticky HDP-HMM を提案している。

## 2.4 sticky HDP-HMM

Fox ら<sup>12)</sup> らは HDP-HMM の自己遷移確率にバイアスを付加したスティッキー階層ディリクレ過程隠れマルコフモデル (sticky Hierarchical Dirichlet Process - Hidden Markov Model: sticky HDP-HMM) を提案している<sup>12)</sup>。自己遷移確率を大きくすることで潜在状態の過剰遷移を抑えることができるため、動作の連続性を仮定するモデリングを効率的に行う事が可能になる。また、Fox らは Blocked Gibbs Sampler による定式化をしめし、より高速にサンプリングを行うことを可能にしている。詳細な議論については文献<sup>12)</sup>を参照されたい。

以下に、sticky HDP-HMM のアルゴリズムを示す。

まず、その時点での状態依存の遷移確率の集合を  $\pi^{(n-1)}$ 、グローバルな遷移確率を  $\beta^{(n-1)}$ 、及び出力分布のパラメータを  $\theta^{(n-1)}$  とする。下記を繰り返す。

1.  $\pi = \pi^{(n-1)}$ 、及び  $\theta = \theta^{(n-1)}$  とする。後ろ向けに計算し、メッセージ  $m_{t,t-1}(k)$  を計算する。

- (a) 全ての  $k \in \{1, \dots, L\}$  を以下に初期化する。

$$m_{T+1,T}(k) = 1$$

- (b) 全ての  $t \in \{T-1, \dots, 1\}$  及び  $k \in \{1, \dots, L\}$  について

$$m_{t,t-1}(k) = \sum_{j=1}^L \pi_k(j) N(y_t; \mu_j, \Sigma_j) m_{t+1,t}(k)$$

2. 時間を順方向に順に状態  $z_{1:T}$  の割り当てをサンプリングしていく。はじめ全ての  $(j, k) \in \{1, \dots, L\}^2$  について  $n_{jk} = 0$  および  $Y_k = \emptyset$

- (a) 全ての  $k \in \{1, \dots, L\}$  について以下の確率を計算する。

$$f_k(y_t) = \pi_{z_{t-1}}(k) N(y_t; \mu_k, \Sigma_k) m_{t+1,t}(k)$$

- (b) 状態の割り当て  $z_t$  をサンプリングする。

$$z_t \sim \sum_{k=1}^L f_k(y_t) \delta(z_t, k)$$

- (c)  $n_{z_{t-1}z_t}$  に 1 加算し、 $y_t$  を、新たな割り当てである  $z_t = k$  のキャッシュ統計に追加する。

$$Y_k \leftarrow Y_k \oplus y_t$$

3. 補助変数  $m, w, \bar{m}$  をサンプリングする。

- (a) 全ての  $(j, k) \in \{1, \dots, L\}^2$  について、 $m_{jk} = 0$  および  $n = 0$  とする。レストラン  $j$  で皿  $k$  を食べている客  $i = 1, \dots, n_{jk}$  についてサンプリングする。

$$x \sim \text{Ber}\left(\frac{\alpha\beta_k + \kappa\delta(j, k)}{n + \alpha\beta_k + \kappa\delta(j, k)}\right)$$

同時に  $n$  をインクリメントし、また  $x = 1$  であれば  $m_{jk}$  をインクリメントする。

- (b) 全ての  $j \in \{1, \dots, K\}$  についてレストラン  $j$  における上書き変数の数をサンプリングする。

$$w_i \sim \text{Binominal}\left(m_{jj}, \frac{\rho}{\rho + \beta_j(1 - \rho)}\right)$$

皿  $k$  を考慮し レストラン  $j$  の informative テーブルの数を設定する。

$$\bar{m}_{jk} = \begin{cases} m_{jk} & j \neq k \\ m_{jj} - w_j & j = k \end{cases} \quad (3)$$

4. グローバル遷移確率分布  $\beta$  をサンプリングする。

$$\beta \sim \text{Dir}(\gamma/L + \bar{m}_{\cdot 1}, \dots, \gamma/L + \bar{m}_{\cdot L})$$

5. 全ての  $k \in \{1, \dots, L\}$  についてサンプリングし

た状態の割り当てに基づいて、新しい遷移確率と出力分布のパラメータをサンプリングする．

$$\begin{aligned}\pi_k &\sim \text{Dir}(\alpha\beta_1 + n_{k1}, \dots, \\ &\quad \alpha\beta_k + \kappa + n_{kk}, \dots, \alpha\beta_L + n_{kL}) \\ \theta_k &\sim p(\theta|\lambda, Y_k)\end{aligned}$$

6.  $\pi^{(n)} = \pi, \beta^{(n)} = \beta, \theta^{(n)} = \theta$  に固定する．

7. (オプション) ハイパーパラメータ  $\gamma, \alpha, \kappa$  をサンプリングする．詳細は文献<sup>12)</sup> 参照．

隠れ状態に対応し変数を出力する確率分布が多次元ガウス分布であるとする、出力分布のパラメータは  $\theta_k = \{\mu_k, \Sigma_k\}$  となる．この時、平均  $\mu_k$  の事前分布をガウス分布  $N(\mu_0, \Sigma_0)$ 、分散共分散行列  $\Sigma_k$  の事前分布を逆ウィシャート分布  $IW(\nu, \Delta)$  とすると、それらのサンプリングの式は以下ようになる、

$$\Sigma_k | \mu_k \sim IW(\bar{\nu}_k, \bar{\Delta}_k, \bar{\nu}_k) \quad (4)$$

$$\mu_k | \Sigma_k \sim N(\bar{\mu}_k, \bar{\Sigma}_k) \quad (5)$$

ここで

$$\bar{\nu}_k = \nu + |Y_k| \quad (6)$$

$$\bar{\nu}_k \bar{\Delta}_k = \nu \Delta + \sum_{y_t \in Y_k} (y_t - \mu_k)(y_t - \mu_k)' \quad (7)$$

$$\bar{\Sigma}_k = (\Sigma_0^{-1} + |Y_k| \Sigma_k^{-1})^{-1} \quad (8)$$

$$\bar{\mu}_k = \bar{\Sigma}_k (\Sigma_0^{-1} \mu_0 + \Sigma_k \sum_{y_t \in Y_k} y_t) \quad (9)$$

である．また、 $Y_k = \{y_t | z_t = k\}$  は観測値の集合であり、 $|Y_k|$  はその集合の濃度を表す．このサンプリングはマルコフ連鎖モンテカルロ法であるため、一度のサンプリングでは必ずしも適切な標本が得られるとは限らず、次のステージに進む前に、 $\mu_k, \Sigma_k$  のサンプリングを交互に数度行うべきである．

### 3 実験

sticky HDP-HMM の非分節動作模倣学習における有効性を検証するために、実験を行った．モーションキャプチャによって計測された多様な単位動作の連続により形成される時系列データを単位動作に分節化することなくモデリングを行った．実験では被験者に複数の対象物を操作させ、それぞれの対象物に応じた動作を行わせることで複数の単位動作を提示させた．モーションキャプチャにより計測したデータに対して Blocked Gibbs Sampler により学習した sticky HDP-HMM と Baum-Welch アルゴリズムにより学習した HMM の結果と比較した．

#### 3.1 モーションキャプチャによる計測と低次元化

人間の上半身の動作時系列をモーションキャプチャにより計測する．計測されたデータを  $X = [x(1)x(2)\dots x(T)]$  とする． $x(t)t = 1, 2, \dots, T$  は各時刻  $t$  に計測された  $D$  次元の列ベクトルであり、 $T$  は時系列長さである．計測されたデータは高次元であるため計算コストの削減を行うことが望ましい．また、人

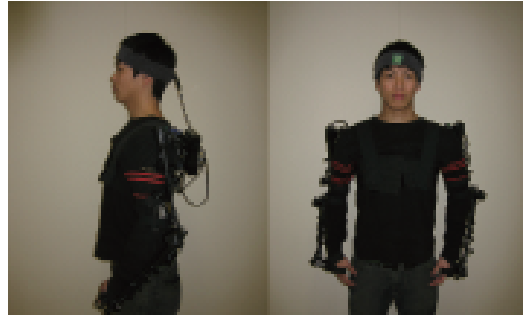


Fig. 3: Motion capture

間の特徴的な動作はその動作の一部の部分空間であると考えられるため、単位動作を抽出するために本質的でない次元の動作を削除する<sup>17)</sup>．そこで、計測されたデータに対して特異値分解を用いて低次元化を行う． $X$  に対して特異値分解を行うと  $X = U\Sigma V^T$  となる．ただし、 $\Sigma$  は  $D \times T$  行列 ( $\Sigma = [\omega_{ij}], \omega_{ij} = \sigma_i \delta_{ij} (1 \leq i \leq D, 1 \leq j \leq T)$ )、 $U$  は  $D \times D$  行列、 $V$  は  $T \times T$  行列である．また、 $\tau$  は転置を表す．このとき、 $\Sigma$  を  $K < D$  として  $\Sigma_{(K)} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_K)$  によって置き換え、 $U, V^T$  をそれぞれはじめの  $K$  列、 $K$  行のみに制限した行列  $U_{(K)}, V_{(K)}$  により作られる行列

$$X = U_{(K)} \Sigma_{(K)} V_{(K)}^T \quad (10)$$

は rank が  $K$  の行列による  $X$  のフロベニウスノルムの意味での最良近似となることが知られている．よって、上記の操作により、時系列データとしては二乗誤差が最小となる低次元時系列による近似となっている．なお、計測データに対して累積率寄与率が 70% までを有効な次元とし、本実験では結果的に 6 次元まで低次元化を行った．

#### 3.2 実験条件

抽出を行う対象を人間の上半身動作とし、機械式モーションキャプチャである Meta Motion 社製の Gypsy5 Torso を用いて人間の各関節角を計測した．Fig. 3 にモーションキャプチャの外観を示す．自由度は頭、首、両肩、両鎖骨、腰、両肘、両手首、胸の 3 次元オイラー角の 36 次元である．実験では、1 セッション 20[sec] とし、3 セッションの動作系列をモーションキャプチャによって記録した．フレームレートは 60[Hz] であり、1 セッションは 1200[step] で構成される．被験者は成人男性である．

被験者には操作対象物として”ボール”、”指示棒”、”Aibo”を順に与え、それぞれを操作するための動作を行うように求めた．これらに対して被験者が操作を行う順序は一定ではなく、被験者が対象物を自由に選んで操作を行った．これらの時系列データは事前に分節化されず、また、操作の間には自然な持ち替え動作やあそびの動作が挿入されていた．Fig. 4 に計測された動作時系列データに含まれていた単位動作を示す．

sticky HDP-HMM のパラメータとしては  $\alpha = 0.1, \gamma = 0.1, \kappa = 1.0$  とした．また、観測データの平均値と分散を求めそれらを出力分布の中心のハイパーパラメータ  $\mu_0, \Sigma_0$  とした．また、分散のハイパーパラメータとして自由度  $9 = (\text{出力次元}) + 3$  及び  $\Delta = dI$

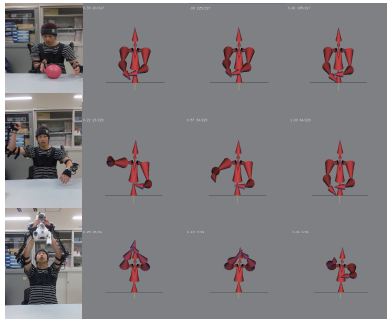


Fig. 4: Target Objects and elemental motion for them

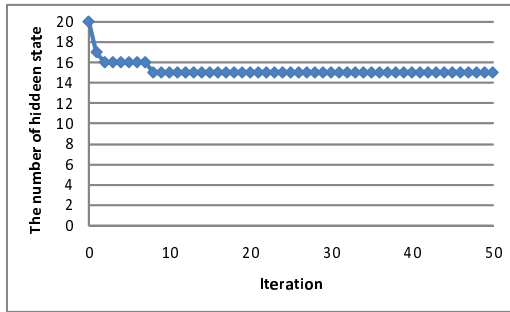


Fig. 5: Transition of the number of hidden state

とし  $d = 1.0 \times 10^{-4}$  とした．後の実験で推定結果の  $d$  に対する依存性を観察する．また，HMM の学習では分散共分散行列が零行列に近づくとき計算が不安定化することがあり正則化項として分散共分散行列に  $\epsilon I$  を付加した． $I$  は単位行列であり， $\epsilon = 1.0 \times 10^{-4}$  は微量である．

### 3.3 実験結果

sticky HDP-HMM による動作時系列のモデル化の結果について述べる．Fig. 5,6 に潜在状態数の推移と潜在状態の事後確率の変化を示す．

潜在状態数は 15 にほぼ収束していった．Fig. 6 からわかるように，基本的にギブスサンプリングであるために，収束することなく，振動している．しかしながら，初期値は本来の事後確率分布において確率の低い部分にあり，そこから徐々に事後確率の高いパラメータのサンプリングとなっている事がわかる．全体的に上昇傾向があるが，サンプリングであるので，前半でも高い値は出力される．この事例の場合，最大

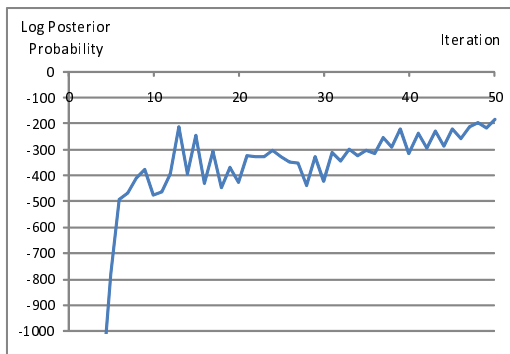


Fig. 6: Log Posterior Probability for estimation of hidden state

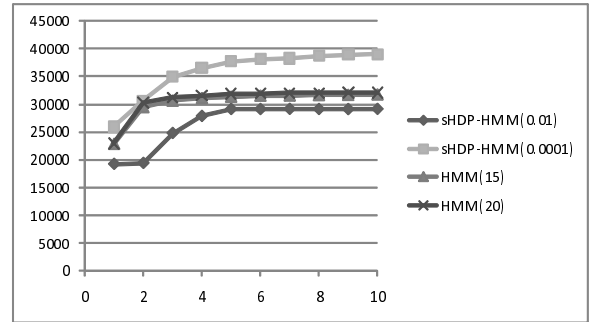


Fig. 7: Comparison of log-likelihood between sticky HDP-HMM and HMM

値を追うと 12 回目で大きな値が出力されている．

次に，通常の HMM の学習との比較を示す．本来，sticky HDP-HMM と Baum-Welch アルゴリズムによる HMM は評価関数が異なるため，公平な比較を行うのはこのような課題の場合困難であるが，仮に対数尤度で比較したグラフを 7 に示す．これは，尤度の最大値を 1 として正規化したグラフである．この時，HMM では状態数を変化させ，sticky HDP-HMM では，各ガウス分布の事前分布である逆ウィシャート分布のパラメータ  $d$  を変化させた．グラフの凡例の括弧内がそれぞれの値を示している．sticky HDP-HMM が尤度の点でも優秀な結果を示していることがわかる．しかし，本来，HMM は尤度最大化を基準としているために，sticky HDP-HMM より低い値となるのは奇異である．これについては二点の影響が考えられる．一つ目は，正則化項の存在である．今回 HMM の学習では分散共分散行列の逆行列計算の不安定化を避けるため，正則化項を導入している．しかし，正則化項は真に尤度を最大化する分散共分散行列の推定値をバイアスする．このため尤度が低下する<sup>7</sup>．二つ目は，局所解の存在であり，Baum-Welch アルゴリズムが局所探索であるために，なんらかの局所解に陥っている可能性がある．これに対して，ギブスサンプリングを用いる sticky HDP-HMM は基本的に大域最適性が保証されており，安定した学習が期待できる．

しかし，sticky HDP-HMM が何の設計も要さないわけではない．次に，sticky HDP-HMM における出力分布となるガウス分布の事前分布である，逆ウィシャート分布のハイパーパラメータ  $\Delta$  を変化させて，隠れ状態の数がどう変化するかを観察した．Fig.8 に横軸に逆ウィシャート分布のハイパーパラメータ  $\Delta = dI$  の  $d$  をとり，縦軸に推定された隠れ状態数をとっている．観測変数のオーダーは大体  $1.0 \times 10^{-2}$  程度であったので， $d = 1.0 \times 10^{-4}$  あたりがオーダー的には妥当なものである．今回は，大局的な関係を見るために大きく値を変化させた． $d$  の値を大きくすると，推定される隠れ状態数が減ることがわかる．これは，事前分布の影響が観測変数よりも相対的に大きくなり，より分散の大きな出力分布により観測変数を説明する方向に全体が動いたことを示している．その結果，自動的にガウス分

<sup>7</sup>これは，分散共分散行列が分散の大きな事前分布を持つことと似た効果を持つ．正則化項をこれ以上小さくすると，本実験では計算の不安定化が見られた．これは数値計算を工夫することで回避可能な面もあるが，この問題は混合分布の特異性と関連しており理論上の問題点でもある．

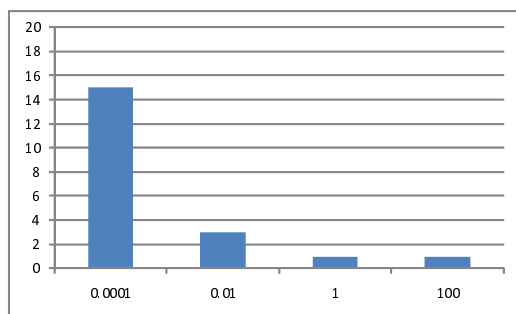


Fig. 8: The number of hidden state depending on the hyper parameter  $\Delta = dI$  of Inverse Wishart distribution

布の数が減ることになる。つまり，sticky HDP-HMM では，個別の出力分布の大きさがハイパーパラメータを通じて緩やかに決定され，その上で従属的に隠れ状態数が決定されるという，従来の HMM とは逆の論理が存在している。もちろん，ベイズモデルとして統合されているので，ディリクレ過程の集中度パラメータなどを通じたトップダウンの制約も存在し，これらボトムアップ，トップダウンの組み合わせで柔軟に隠れ状態数，出力分布が決定するモデルであることがわかる<sup>8</sup>。

#### 4 まとめ

本研究では非分節動作模倣学習のために，ノンパラメトリックベイズモデルを用いる事について検討した。具体的にはスティッキー階層ディリクレ過程隠れマルコフモデル (sticky HDP-HMM) を導入し，その有効性について確認した。HMM では，従来は実験者によって決定されていた潜在状態数を学習器自体が自動的に決定し，さらに実データに対して，大域最適解に近い解を効率的に探索しうる可能性が示された。また，ハイパーパラメータを変化させ実験を行うことで，出力分布からの影響がボトムアップに現れ，隠れ状態数が決定することを示した。

非分節動作模倣学習は，sticky HDP-HMM によるモデリングだけでは，完成せずに，符号化された隠れ状態列をさらに単語列へと分節化するプロセスが必要になる。谷口らは符号化・モデリングについては各出力分布が線形モデルである HMM である SARM を用い尤度最大化基準の Baum-Welch アルゴリズムで学習させ，その結果を MDL 基準で局所探索することで単語抽出を行っているが<sup>5)</sup>，これら二つのプロセスを統合することが課題である。そのためにも統一された確率モデルによって，この二つのプロセスを表現する事が重要である。持橋らはノンパラメトリックベイズモデルにより書き言葉の教師無し形態素解析を実現しており，ノンパラメトリックベイズモデルは上位下位のプロセスを統合する柔軟性を持っていると期待される<sup>18)</sup>。ノンパラメトリックベイズモデルに基づいて，単語抽出と動作時系列の符号化を同時に行うモデルを開発することが今後の課題である。

#### 参考文献

- 1) 明和政子. 心が芽ばえるとき コミュニケーションの誕生と進化 (叢書コムニス). NTT 出版, 2006.
- 2) 谷口忠大. コミュニケーションするロボットは創れるか - 記号創発システムへの構成論的アプローチ (叢書コム

<sup>8</sup>これらの議論は sticky HDP-HMM にとどまらずディリクレ混合過程に類したクラス全体の特性である。

ニス 13). NTT 出版, 3 2010.

- 3) 谷口忠大, 岩橋直人, 中西弘門, 西川郁子. ヒューマン・ロボットインタラクションを通じた役割反転模倣に基づく実時間応答戦略獲得. 第 23 回人工知能学会全国大会 in CD-ROM, 2009.
- 4) T. Taniguchi, H. Nakanishi, and N. Iwahashi. Simultaneous estimation of role and response strategy in human-robot role-reversal imitation learning. In *The 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, 2010.
- 5) 谷口忠大, 岩橋直人. 複数予測モデル遷移の n-gram 統計に基づく非分節運動系列からの模倣学習手法. 知能と情報 (日本知能情報フレンジ学会論文誌), Vol. 21(6), pp. 1143–1154, 2009.
- 6) J. Tani, M. Ito, and Y. Sugita. Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Networks*, Vol. 17, No. 8-9, pp. 1273–1289, 2004.
- 7) 藤田雅博, 下村秀樹 (編). 発達する知能 - 知能を形作る相互作用 (インテリジェンス・ダイナミクス). シュプリンガー・ジャパン, 2008.
- 8) M. Okada, D. Nakamura, and Y. Nakamura. Self-organizing Symbol Acquisition and Motion Generation based on Dynamics-based Information Processing System. In *Proc. of the second International Workshop on Man-Machine Symbiotic Systems*, pp. 219–229, 2004.
- 9) T. Taniguchi, N. Iwahashi, K. Sugiura, and T. Sawaragi. Constructive approach to role-reversal imitation through unsegmented interactions. *Journal ref: Journal of Robotics and Mechatronics*, Vol. 20, No. 4, pp. 567–577, 2008.
- 10) 稲島哲也, 谷江博昭, 中村仁彦. 統計的行動認識・生成モデルの幾何空間における原始シンボルの発達と操作. 電子情報通信学会, Vol. 103, No. 392, pp. 37–42, 2003.
- 11) 上田修功, 山田武士. ノンパラメトリックベイズモデル. 応用数理, pp. 196–214, 2007.
- 12) 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, November 2007.
- 13) T.S. Ferguson. A Bayesian analysis of some nonparametric problems. *The annals of statistics*, Vol. 1, No. 2, pp. 209–230, 1973.
- 14) J. Sethuraman. A constructive definition of Dirichlet priors. *Statistica Sinica*, Vol. 4, No. 2, pp. 639–650, 1994.
- 15) Y.W. Teh and M.I. Jordan. Hierarchical Bayesian nonparametric models with applications. *Bayesian Nonparametrics*, p. 158, 2009.
- 16) M.J. Beal, Z. Ghahramani, and C.E. Rasmussen. The infinite hidden Markov model. *Advances in Neural Information Processing Systems*, Vol. 1, pp. 577–584, 2002.
- 17) T. Taniguchi and N. Iwahashi. Imitation learning from unsegmented human motion using switching autoregressive model and singular vector decomposition. In *Advances in Neuro-Information Processing*, No. 5506, pp. 953–961. Springer, 2009.
- 18) 持橋大地, 山田武士, 上田修功. ベイズ階層言語モデルによる教師なし形態素解析. 情報処理学会研究報告, 2009.