

ディリクレ過程混合モデルを用いた移動ロボットの概念形成

Acquiring Internal Models by a Mobile Robot using Dirichlet Process Mixture

山下 純平*¹ 谷口 忠大*¹
Jumpei Yamashita Tadahiro Taniguchi*¹立命館大学情報理工学部

College of Information Science and Engineering, Ritsumeikan University

There is increasing interest in noble studies about autonomos robots. Traditional autonomous robots do not work without implementing their internal models. Therefore, it is important to design an autonomous robot which is able to to segment environment and conceptualize its own internal models by itself. This paper assumes that generation-process which is about data of robot's motors and sensors is Dirichlet Process Mixture. This enables an autonomous mobile robot embedded with several sensors and motors to acquire some internal models through its interaction with the environment. We conducted a simple simulation experiment in which a mobile robot categorizes its sensorimotor information into three internal models corresponding to three situations in a corridor despite the number of internal models had not been given.

1. はじめに

自律ロボットは自らのセンサ・モータ系に閉じた環世界 [1] の中で、環境情報を分節化し意味づけせねばならない。Taniらは Recurrent Neural Network(RNN) を並列かつ階層的に配置した the mixture of RNN experts を持たせた移動ロボットを操作し、自律的にロボットがセンサ・モータ系の情報の組織化をとおして複数の予測モデルを獲得し、それに基づいて環境の分節化を行わせていた [2]。Wolpert らによる MOSAIC は人間の中枢神経系において獲得されるとする、多重順逆モデルの計算論であり、予測誤差から計算される責任信号により各予測モジュールが重みづけられ、学習および予測を行うことでセンサ・モータ情報の分節化を行う [3]。MOSAIC では各モジュールの線形性を仮定するがこの線形性の仮定は様々な良い性質を持つ面もある。Doya らは MOSAIC をもとにしたモジュール型の強化学習モデル (MMRL: Multiple model-based reinforcement learning) を提案しているが、ここでは予測モデルが線形性を持つということと報酬関数が二次形式であるという仮定から、局所的な領域における最適制御器を明示的に導出することが出来ている [4]。鮫島らはこの手法に基づきながら模倣学習の仕組みを提案している [5]。これらの手法は、各モジュールの分担にベイズ推論を用い、勾配法を用いて各モジュールの学習則を導出するものが主であるが、ベイズ的な生成モデルとして構築することで、性質のよい線形のハイブリッドダイナミカルシステムについては EM アルゴリズムや MCMC を用いた学習則の導出も可能になる [6]。谷口らは、この Switching AR model を用いて人間の動作を線形予測器の切替えに基づいて分節化したのちに、その定型的な符号遷移を結合し模倣学習する手法を提案している [7]。しかし、これらの手法では学習前にロボットが分節化する環境情報の種類、つまり概念の数を決定しなければならない。この問題を解決するために、谷口は仮説検定に基づいたシエマモデルと呼ばれるモジュール型の学習器を開発し、一連の研究を行った [8, 9]。これにより自律ロボットはオンラインで累積的にモジュールを獲

得し、概念形成を行うことが出来た。しかしながら、その挙動に対する理論的解析は困難で、学習結果も安定的に得られるとは言い難いという問題があった。

これに対して近年、ベイズ理論に基づきながらオンラインでのモジュール数の適応的な調整が可能なノンパラメトリックベイズモデルによるアプローチが注目されている [10, 11]。本稿では局所的な線形予測器の数を適応的に配置することの出来るディリクレ過程混合モデル (DPM: Dirichlet Process Mixture) を移動ロボットに持たせ、Tani らの課題 [2] に類した迷路の中の移動を、人間が操作することによってシミュレーション環境上で行わせ、概念の数*¹を自律ロボットがボトムアップに獲得出来ることを示す。

2. 手法

本稿ではディリクレ過程混合モデルおよびこれを用いた可変数個の局所線形線形予測器について説明する。

2.1 ディリクレ過程混合モデル

ディリクレ過程混合モデルはディリクレ過程 (DP: Dirichlet Process) を各出力モジュールの選択確率の事前分布に利用した混合モデルである。ディリクレ過程は無次元のディリクレ分布であり、原理的には無限次元の多項分布を生成する分布である。

ディリクレ過程は、基底分布 G_0 と正の集中度パラメータ γ で定義される。ディリクレ過程混合モデルのデータ生成過程は以下である。

$$(1) \quad G \sim DP(\gamma, G_0) \quad (1)$$

$$(2) \quad \theta_i | G \sim G, \text{ for } i = 1, \dots, n \quad (2)$$

$$(3) \quad x_i \sim p(x|\theta_i), \text{ for } i = 1, \dots, n \quad (3)$$

まず、パラメータ θ の事前分布の事前分布である基底分布 G_0 から、DP によってパラメータの事前分布 G が生成される。これは $\theta \in \Theta$ を変数に持つ離散分布である。データが生成される際には、そのデータ毎に分布 G から θ_i を生成し出力分布の

連絡先: 山下純平, 立命館大学院理工学研究科, 創発システム研究室, 滋賀県草津市野路東 1-1-1, 077-561-5829, E-mail: yamashita@em.ci.ritsumeikan.ac.jp

*¹ つまり、局所線形予測器の数

パラメータとしている．DP の具体的な実装方法として Stick-breaking process(SBP) や Chinese restaurant process(CRP) を用いたものがある [12]．本稿では CRP に基づきモデル化し Gibbs Sampling を用いたクラスタリングによりモデリングを行う．

2.2 Chinese Restaurant Process

CRP は前に述べた DP を Gibbs Sampling を用いて実装する方法を与えるものである．CRP では既に生成されたデータの所属数が多いクラスほど，その所属数に比例して新たに生成されるデータがそのクラスに分類される確率が高まる．また，ある一定の確率で新たなクラスを作りこれに分類する．この際，基底分布から新たに作られたクラスに対応するパラメータが生成される．既存のクラスまたは新規のクラスが採用される確率を次に式として示す．

$$P(z_i = k | z_{1:i-1}) = \begin{cases} \frac{m_k}{i-1+\gamma} & m_k > 0 \\ \frac{\gamma}{i-1+\gamma} & k \text{ は新規クラス} \end{cases} \quad (4)$$

$z_{1:i-1}$ は既に生成されているデータそれぞれが所属するクラスを表した数値列である． z_i は次に決定するクラスであり，次に生成されるデータのクラスである． m_k は，ある既存クラス k に属するデータ数である． γ は DP におけるパラメータである．

2.3 Gibbs Sampling によるクラスタリング

$P(z_{1:n} | x_{1:n})$ を最大化する際データの生成過程に DPM を仮定し，DP の実装に CRP を用いると Gibbs Sampling を使って近似的かつ確率的に値の大きな $P(Z_{1:n} | X_{1:n})$ を求めることができる．以下にそのための式を示す．

$$P(z_i = k | z_{-i}, x_i, x_{-i}) = \begin{cases} \frac{m_{-i,k}}{n-1+\gamma} \int p(x_i | \theta_{(k)}) P(\theta_{(k)} | x_{-i}) d\theta_{(k)} \\ \frac{\gamma}{n-1+\gamma} \int p(x_i | \theta_i) G_0(\theta_i) d\theta_i \end{cases} \quad (5)$$

ただし，

$$P(\theta_{(k)} | x_{-i}) = \frac{\prod_{s: z_s = k, s \neq i} p(x_s | \theta_{(k)}) G_0(\theta_{(k)})}{\int \prod_{s: z_s = k, s \neq i} p(x_s | \theta) G_0(\theta) d\theta} \quad (6)$$

このアルゴリズムでは，はじめにクラス数，データそれぞれのクラスを任意に設定する．次に，上記の式を用いて逐次的にデータそれぞれの新しいクラスを決めることで，事後確率 $P(Z_{1:n} | X_{1:n})$ をサンプリングする．これにより入力データをクラスタリングすることが出来る．

2.4 可変個の局所線形予測器の獲得

上記のディリクレ過程混合モデルを用いる事で複数の局所的な線形予測器を移動ロボットに獲得させることが出来る．一般に出力分布に線形予測器を用いることも可能であるが，入力空間全てにわたり存在するがゆえの汎化能力が逆に望ましくない結果を生むことがあるため，複数予測器の学習では，入力分布に局所性を持たせる事がある．Doya らの MMRL でガウス分布を入力空間のバイアスとして導入している．線形モデルの出力分布の出力ノイズにガウス分布を仮定し，入力分布をガウス分布とした場合には，同時分布はガウス分布となる．故に，出力分布をガウス分布とした DPM により，前後の時刻にまたがったセンサ・モータ情報の同時分布をモデリングすることにより，モジュール型の局所線形予測器を複数配置し推定した場合と同等の結果を得ることが出来ると考えられる．

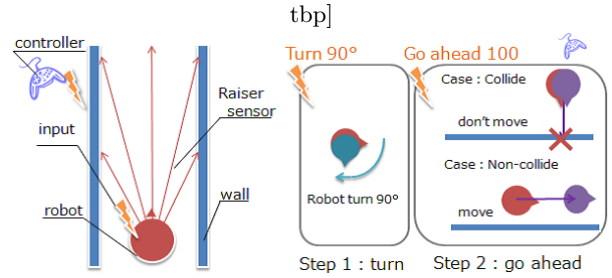


図 1: ロボットの構成と動作の概観

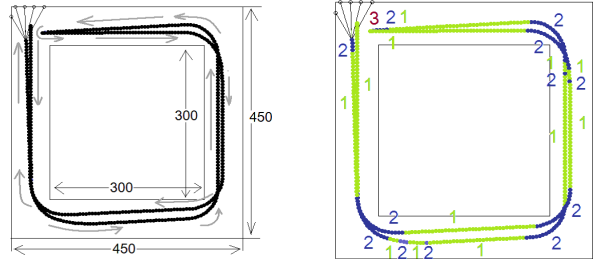


図 2: 実験環境とロボットの軌道 図 3: 得られた予測モデルと軌道の対応関係

3. 実験

DPM を複数内部モデルの生成器として持つ移動ロボットを操作し，環境の予測モデルを形成させる実験を行った．

3.1 実験条件

3.1.1 移動ロボット

移動ロボットの構成の概観は図 1 のようになっている．移動ロボットはセンサー・モータ系を持ち教示者がコントローラのアナログスティックを介して操作を行う．モータ出力は前進距離と回転角度を入力とし 2 次元ベクトル a_t で表される．また，センサ入力には 5 つの距離センサを持つとして，有効範囲は d_{max} とし，各々のセンサが障害物までの距離 d に対して $s = 1 - d/d_{max}$ を返す．これにより 5 次元のセンサ入力ベクトル s_t を得る．ここで $d_{max} = 100.0$ とした．移動する際に，旋回は必ず行われる (図右方左枠) が，直進はその結果壁にぶつかるとすればその場に留まる (図右方右枠)．

予測器の学習に教示学習時のロボットの走行から与えられる時系列データを用いる．それはある時刻 t のセンサ・モータ情報 s_t, a_t とその次の時刻センサ情報 s_{t+1} の変化量 $\delta s_t = s_{t+1} - s_t$ がセットになったものを対象とする $(\delta s_t, s_t, a_t)^{*2}$ ．個別の予測器に相当する DPM の出力分布をの多次元ガウス分布とし，学習データに対する同時確率分布 $p(\delta s_t, s_t, a_t)$ をモデリングする．ガウス分布の性質から，ある時刻におけるセンサ・モータ情報を条件部とした次時刻のセンサ情報の条件付き確率分布 $p(s_{t+1} | s_t, a_t)$ が得られるため [13]，局所線形予測器を獲得しているのと同様の意味を持つ．

3.1.2 実験環境

今回の実験ではロボットは，正方形の外壁と同じく正方形の内壁に囲まれた通路を移動する．移動軌跡は図に示す通りで，図左上方から下向きに出発し反時計回りに一周し，一周すると U ターンして今度は時計回りに一周するようにした (図 2)．

*2 s_{t+1} と s_t には強い相関があるために，このように差分をとることで，今回はその相関に影響されすぎないようにした．

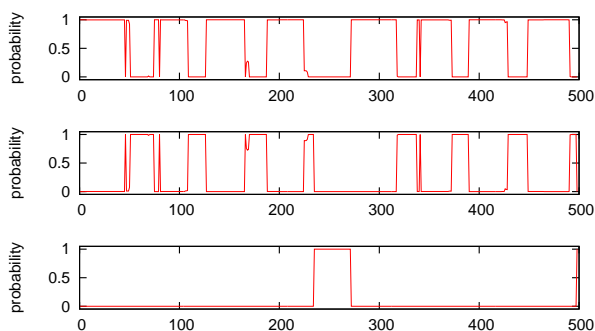


図 4: 各時刻におけるセンサモータ情報が各クラスに所属する確率

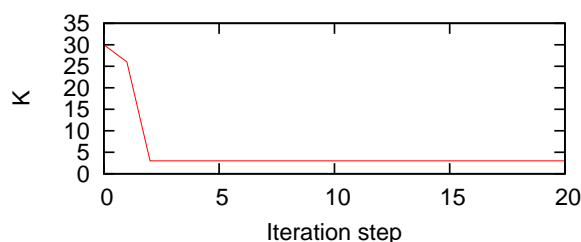


図 5: Gibbs Sampling の繰り返し数に対する内部モデルの数の変化

この環境の定数は図中に示す．この際、モータ出力には入力値には各入力に対して独立に平均 0.0、分散 1.0×10^{-4} の微小なガウスノイズを乗せて移動させた．また、図には学習データに対応した時刻の座標それぞれを中心に、点を打ち軌跡として表している．

基本的な設定としては、DPM の初期クラス数は 30、Gibbs Sampling の繰り返し回数は 20 回、集中度パラメータ $\gamma = 1.0$ とした．また、多次元ガウス分布のパラメータの事前分布については、平均の事前分布には平均 0、分散 1.0 のガウス分布を、分散の事前分布には自由度 12、及び正定行列には単位行列を用いた．

3.2 結果

Gibbs Sampling の繰り返し学習データがクラスターリングされた結果を図 3 に示す．3 つのクラスが生成され (図 5)、図 3 に示すようにデータはそれぞれ 3 つのクラスに分類された．3 つめのクラスは一見分かりづらいが一周を終えた直後の旋回部分で割当てられており、これは図 4 グラフからも見て取れる．

図 4 にクラス毎に時刻のデータが各予測器に割り当てられる確率を表した．上から順に 1 番から 3 番の内部モデルが選択される各時刻での事後確率値を表している．一番下のクラスのグラフから、ちょうど折り返しの 250 あたりでデータがクラス 3 に割り当てられることが見て取れる．また各時刻に置いて先ほど示した学習データ軌跡にあわせたクラス分けの図と一致していることが見て取れる^{*3}．

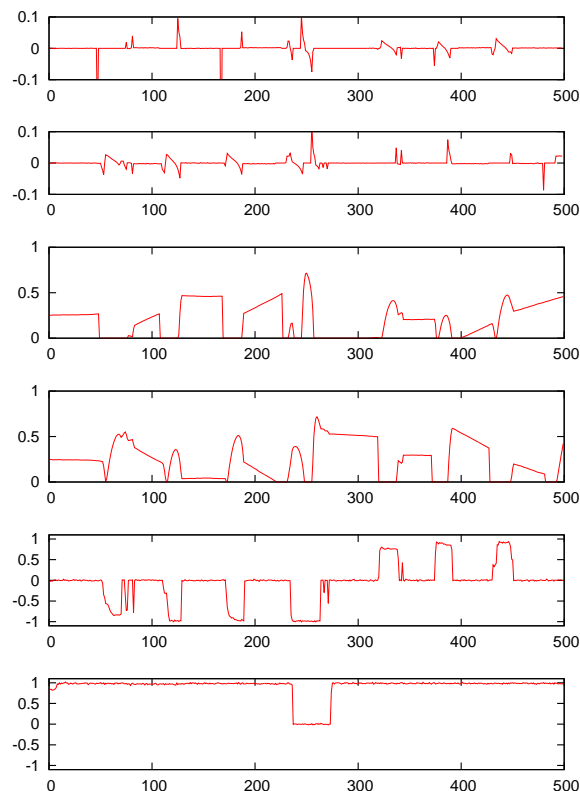


図 6: 各時刻におけるセンサモータ情報の値．上から順に各時刻の $\delta s^1, \delta s^5, s^1, s^5, a^1, a^2$ の順に示す．

4. 考察

結果から見られるように、内部モデル 1~3 それぞれのクラスは直線的に移動、角を曲がる、旋回する状況に対応しており、移動ロボットは環境との相互作用のみを通してボトムアップに複数の予測モデルをその数と共に学習することができた．

ここで何が理由で、それぞれの予測モデルにまとまったのかを検討する．内部モデルは予測モデルであり、自らの行為に対するセンサ値の変化を基にクラスが分かれるが、それを同時分布を通して推定しているため、モータ出力値 a_t にも影響を受ける．図 6 に特徴的なセンサモータ値を各ステップについて示す． s^1, s^5 はそれぞれ左右両端に位置する距離センサである．クラス 3 は明らかに、 a^2 が 0 に近づいている状態に対応しており、移動ロボットが前進移動をしていない場合に相当する．これは、モータ値の違いも強く効いて一つのクラスを形成しているものと考えられる． $(\delta s_t, s_t, a_t)$ のセンサモータ情報を主成分分析で二次元に射影した図を図 7 に示す．各成分の寄与を見たところ横軸 (第一主軸) は主に前進行動 a^2 が負に、また縦軸 (第二主軸) には旋回行動 a^1 が正に寄与していた

一方で各コーナーは全てクラス 2 に分類されているが、その際の旋回行動のモータ値 a^1 は正負が逆になるゆえに、第二種軸に大きなばらつきが生じる．また第二種軸にはセンサ変化 δs が全成分正に寄与しており、旋回時にセンサ値に大きな揺動が起きることが現われている．順方向ダイナミクスとして直進時の定常的な状態がクラス 1 に、旋回時のセンサモータ値

*3 本来、各予測器へ割り当てられる確率は、ノンパラメトリックに推定しようが、ここでは各ガウス分布のパラメータを最尤推定して、求めたガウス分布から各時刻のデータがあるクラスであり得る確率をグラフにしている．

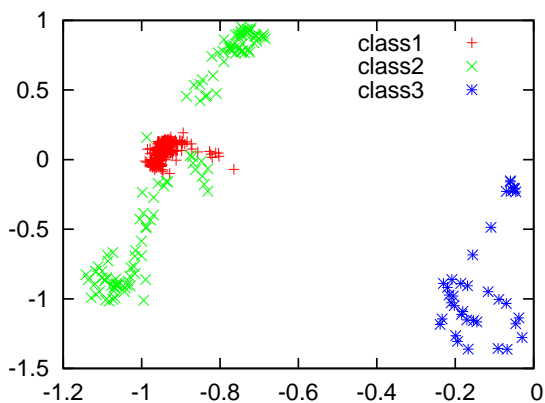


図 7: センサモータ情報の主成分分析による 2 次元空間プロット

の揺動がクラス 2 としてまとまったと考えられる。

また、本実験時の行動データに対しては異なる乱数種を用いて 10 回学習を行ったがその全てに対して、3 つのクラスが出来た。しかし、行動データを変え、蛇行しながら往復するデータにしたところ、クラス数が 1 や 2 となる事が観察された。クラス数がデータセットに依存するのは当然であるが、曲がり角と直進路も区別されなくなっており、このような点についての議論はまた必要である。

5. まとめ

本稿では移動ロボットに複数の予測器を自律的に獲得させるために DPM を用いた。実験においては簡単な迷路環境のシミュレーションにおいて、その有効性を調べた。ガウス分布を出力分布とした DPM は各予測器の入力分布と出力ノイズにガウス分布を仮定した局所線形モデルと見なすことが出来る。このようにボトムアップに線形予測器の個数を獲得できれば、ベイズ理論の良さを維持したままに、谷口らのシマモデルのように個数を可変的に獲得する事が可能になる。現在は予測器をボトムアップに獲得するにとどまるが、Doya の MMRL を始め、環境を線形に分節化、記号化することで、様々な運動制御へと接続することも可能である。ロボット自身の移動行動に接続し、自律行動の創発へとつなげるのが今後の課題である。

参考文献

- [1] ユクスキュル, クリサート. 生物から見た世界 (岩波文庫). 岩波書店, 2005.
- [2] J. Tani and S. Nolfi. Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Networks*, Vol. 12, pp. 1131–1141, 1999.
- [3] D.M. Wolpert and M. Kawato. Multiple paired forward and inverse models for motor control. *Neural Networks*, Vol. 11, pp. 1317–1329, 1998.
- [4] K. Doya, et al. Multiple model-based reinforcement learning. *Neural Computation*, Vol. 14, pp. 1347–1369, 2000.

- [5] 鮫島和行, 銅谷賢治, 川人光男. モジュール競合による運動パターンのシンボル化と見まね学習. *電信情報通信学会論文誌 D-II*, Vol. J85-D-II(1), pp. 90–100, 2002.
- [6] K.P. Murphy. Switching Kalman filters. *Technical reports, DEC/ Compaq Cambridge Research Labs*, 1998.
- [7] 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 (日本知能情報ファジィ学会論文誌)*, Vol. 21, No. 6, pp. 1143–1154, 2009.
- [8] 谷口忠大, 榎木哲夫. 身体と環境の相互作用を通じた記号創発: 表象生成の身体依存性についての構成論. *システム・制御・情報学会誌*, Vol. 18, No. 12, pp. 440–449, 2005.
- [9] 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 (叢書コムニス 13). エヌティティ出版, 2010.
- [10] Y.W. Teh and M.I. Jordan. Hierarchical Bayesian non-parametric models with applications. *Bayesian Non-parametrics*, p. 158, 2009.
- [11] 上田修功, 山田武士. ノンパラメトリックベイズモデル. *応用数理学会誌*, Vol. 17, No. 3, pp. 196–214, 2007.
- [12] Y.W. Teh, M.I. Jordan, M.J. Beal, and D.M. Blei. Hierarchical dirichlet processes. *Journal of the American Statistical Association*, Vol. 101, No. 476, pp. 1566–1581, 2006.
- [13] Christopher M. Bishop. *Pattern Recognition And Machine Learning (Information Science and Statistics)*. Springer-Verlag, new edition edition, 2006.