

Coupled GP-HSMM を用いた連続動作の分節化に基づく インタラクションのモデル化

Interaction Modeling Based on Motion Segmentation Using Coupled GP-HSMM

押川 慧^{*1}
Satoru Oshikawa

中村 友昭^{*1}
Tomoaki Nakamura

長井 隆行^{*1}
Takayuki Nagai

岩橋 直人^{*2}
Naoto Iwahashi

船越 孝太郎^{*3}
Funakoshi Kotaro

中野 幹生^{*4}
Mikio Nakano

金子 正秀^{*1}
Masahide Kaneko

^{*1}電気通信大学
The University of Electro-Communications

^{*2}岡山県立大学
Okayama Prefectural University

^{*3}京都大学
Kyoto University

^{*4}(株)ホンダ・リサーチ・インスティテュート・ジャパン
Honda Research Institute Japan Co., Ltd.

In the human community, there are various interactions and humans can learn them by observing them or interacting with others. For realizing robots that can coexist with humans, it is important for robots to be able to learn appropriate interactions in the community. In this paper, we propose the novel model coupled Gaussian process hidden semi-Markov model (Coupled GP-HSMM) that enables robots to learn rules of interaction between two persons by observing it in an unsupervised manner. The continuous motions of the persons are segmented into discrete actions based on GP-HSMM, and relationships between the actions are extracted. Moreover, all corresponding actions are not simultaneously conducted by two persons in actual interaction and coupled GP-HSMM models such lags between actions. We conducted experiments using motion data of interaction games and experimental results showed that coupled GP-HSMM can estimate actions, lags between them and their relationships.

1. はじめに

近年、ロボットは様々な場面で活躍しており、今後インタラクションを行いながら我々の日常生活をサポートするロボットが期待される。この時、より自然なインタラクションが行われるのが望ましいが、人間社会には地域や文化によって様々なインタラクションのルールが存在するため、環境毎に使用されるインタラクションを事前に全て作りこむことは難しい。人間の場合、他者とのインタラクションや他者同士のインタラクションの観察から、インタラクションで交わされる声や動作の関係性を見つけ、インタラクションのルールを学習することが出来る。また、幼児は身近な人のふるまいを見てそれを模倣して遊ぶことがあり、その模倣を通して動作やインタラクションを学習している。ロボットが人間と共存していくには、人間と同様にインタラクションの観察によって学習することが重要であると言える [Billard 08][Berenna 09][Schaal 99]。

ロボットのインタラクションの学習に関する先行研究として文献 [Taniguchi 10] がある。この研究では役割反転模倣を通じて連続的な実時間インタラクションからロボットがインタラクション内の応答戦略を獲得しており、役割を推定することによってルールを学習している。しかし、この研究ではロボットの行動がシミュレーション上で事前に定義されているため、それ以外の行動やルールを学習することはできない。また、文献 [Soltoggio 13] では遅延報酬を考慮したゲームのルール学習を対象として、実際のロボットを用いて実験を行っているが、この研究においても行動が事前に定義されており、それ以外の行動やルールを学習することはできない。このように、ルール学習に関する研究はなされているが、事前に行動が定義されているものが多く、連続動作からインタラクションの行動とルールを同時に学習する研究は十分になされていない。

そこで、本稿では連続動作を行動へと分節し、二者間の行動の相互関係を推定することで、その間で行われているインタラクションを教師なしで学習するモデルを提案する。本稿における連続動作とはインタラクション内の一連の動きのことであり、行動とは連続動作から意味のある単位ごとに分節・分類された単位動作を意味する。提案モデルではインタラクション内

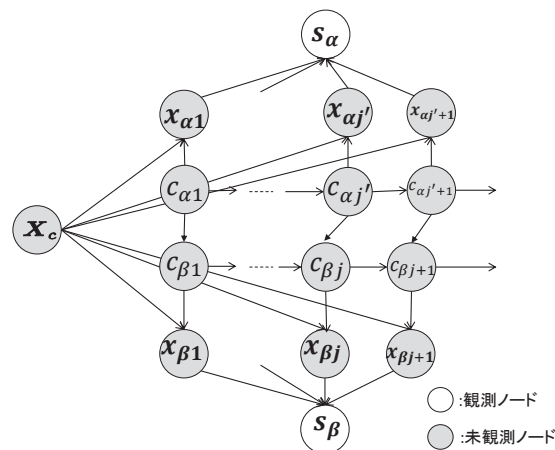


図 1: Coupled GP-HSMM

における連続動作を分節化することで、類似した行動ごとに分類し、分類された行動の相互関係を推定することでインタラクションを学習する。連続動作を教師なしで分節化する手法として、GP-HSMM[Nakamura 17] や Autoplaits[Matsubara 14], BP-HMM[Fox 11] などがあり、本稿ではその中でも分節結果の精度が良いと報告されている GP-HSMM を拡張した Coupled GP-HSMM を提案する。このモデルでは、互いの行動の相互関係を考慮することで、ルールを学習するだけでなく、動作の分節も相補的に学習することができ、分節化の精度の向上が可能となる。

2. Coupled GP-HSMM に基づくインタラクションのモデル化

本稿では、 α と β の二者の連続動作 S_α と S_β を行動 x_α , x_β へと分節し、その行動クラス c_α , c_β の関係性を教師なしで

連絡先: 押川慧, 電気通信大学大学院情報理工学研究所, 東京都調布市調布ヶ丘 1-5-1, s.oshikawa@radish.ee.uec.ac.jp

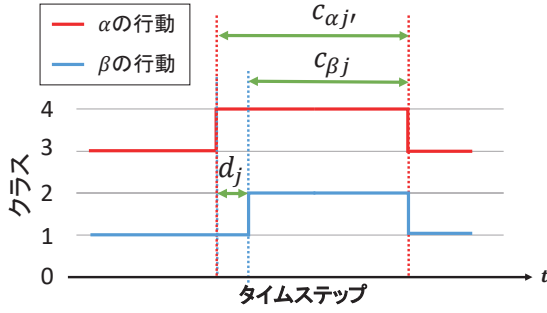


図 2: インタラクションの時間差

学習することで、それらの相互関係をモデル化する。図 1 が提案する Coupled GP-HSMM のグラフィカルモデルであり、灰色が未観測ノードを表している。 $\mathbf{X}_c$ はクラス c のガウス過程のパラメータであり、行動 $\mathbf{x}_{\alpha j'}$, $\mathbf{x}_{\beta j}$ は $\mathbf{X}_{c_{\alpha j'}}$, $\mathbf{X}_{c_{\beta j}}$ をパラメータとするガウス過程によって生成される。GP-HSMM では、 β の j 番目のクラス $c_{\beta j}$ は直前のクラスによってのみ生成されるが、このモデルにおいては β 自身の $j-1$ 番目のクラス $c_{\beta, j-1}$ と、それに対応する α のクラス $c_{\alpha j'}$ によって生成されることを仮定し、以下のような生成過程で表現される。

$$c_{\alpha j'} \sim P(c_{\alpha, j'-1}) \quad (1)$$

$$c_{\beta j} \sim P(c_{\beta, j-1}, c_{\alpha j'}) \quad (2)$$

$$\mathbf{x}_{\alpha j'} \sim \mathcal{GP}(\mathbf{x} | \mathbf{X}_{c_{\alpha j'}}) \quad (3)$$

$$\mathbf{x}_{\beta j} \sim \mathcal{GP}(\mathbf{x} | \mathbf{X}_{c_{\beta j}}) \quad (4)$$

生成されたそれぞれの行動を結合することによって観測系列 $\mathbf{S}_\alpha$, $\mathbf{S}_\beta$ が生成される。この Coupled GP-HSMM は互いの行動が相互に影響するモデルとなっており、実際のインタラクションからこのモデルでインタラクションを学習することで連続動作を行動へと分節しながら互いの行動の相互関係を推定することが出来る。

2.1 ガウス過程

本稿では、行動のタイムステップ i における出力値 x_i をガウス過程回帰で学習することで、取得した時系列データを離散化することなく、連続的な分布として表現する。ガウス過程では同一クラスに属する動作データのタイムステップ i に対する出力 x の組 $(i, \mathbf{x})$ が得られた時、タイムステップ i^{new} における出力 x^{new} の予測分布は式 (13) のガウス分布となる。

$$p(x^{new} | i^{new}, \mathbf{x}, i) \propto N(\mathbf{k}^T \mathbf{C}^{-1} \mathbf{i}, c - \mathbf{k}^T \mathbf{C}^{-1} \mathbf{k}) \quad (5)$$

得られた予測分布を用いることで、 $\mathbf{x}_i$ がクラス c に対応するガウス過程から生成される確率 $\mathcal{GP}(\mathbf{x}_i | \mathbf{X}_c)$ を以下のように求めることができ、この確率を用いることで似た行動を同じクラスに分類することが出来る。

$$\begin{aligned} \mathcal{GP}(\mathbf{x} | \mathbf{X}_c) &= p(x_0 | i, \mathbf{X}_{c,0}, \mathbf{I}) \\ &\quad \times p(x_1 | i, \mathbf{X}_{c,1}, \mathbf{I}) \\ &\quad \times p(x_2 | i, \mathbf{X}_{c,2}, \mathbf{I}) \cdots \end{aligned} \quad (6)$$

2.2 インタラクションにおける行動の時間差

あるインタラクション内で、 α がクラス $c_{\alpha j'}$ の行動を行い、それに対応して β がクラス $c_{\beta j}$ の行動を行う場合、必ずしも同時に行われるわけではなく図 2 のように時間差が生じる場合がある。この時間差には種類があり、一つ目として、実際のインタラクションでは行動の時間差はないが分節の推定誤りで行動に時間差が生じる場合である。提案モデルでは相手との相

互関係を含めて推定を行っているため、このような推定誤りを減らすことが可能である。

もう一つの時間差としては、実際にインタラクションにおける行動がずれている場合である。インタラクションでは、相手と息を合わせて行っている場合、行動がずれてしまい同タイミングで行われることが少ない。このような時間差の場合は、行動の推定誤りではないため、修正することはできない。しかし、相手の行動との関係性からルール抽出を行う場合、このような時間差を無視してしまうと、タイミングがずれた行動同士の関係性を推定できなくなってしまう。そこで、提案モデルではこの時間差 d_j がガウス分布に従うと仮定し、教師なしで推定する。

$$p(d_j) = \mathcal{N}(d_j | 0, \sigma^2) \quad (7)$$

2.3 学習

提案モデルでは、取得した観測系列 $\mathbf{S}_\alpha$, $\mathbf{S}_\beta$ が与えられた時、これを分割して得られる行動 $\mathbf{x}_\alpha = x_{\alpha 1} x_{\alpha 2} \cdots x_{\alpha N}$, $\mathbf{x}_\beta = x_{\beta 1} x_{\beta 2} \cdots x_{\beta N}$ が生成される確率 $P(\mathbf{x}_\alpha, \mathbf{x}_\beta | \mathbf{S}_\alpha, \mathbf{S}_\beta, \Theta)$ を最大化するパラメータを推定する。

$$\arg\max_{\Theta} = P(\mathbf{x}_\alpha, \mathbf{x}_\beta | \mathbf{S}_\alpha, \mathbf{S}_\beta, \Theta) \quad (8)$$

しかし、教師なしで行動へと分節化を行う場合、行動の分節位置もその行動のクラスも未知であり、その双方を推定しなければならない。そのため、モデルの学習には一つの観測系列内の行動とそのクラスをまとめてサンプリングする Blocked Gibbs Sampler を用いる。

2.3.1 Blocked Gibbs Sampler

提案モデルでは、観測系列内の行動とそのクラス、行動の時間差をサンプリングすることで推定する。まず、全ての観測系列をランダムに分節・分類することで初期化する。次に、ある観測系列 $\mathbf{S}_{\beta n}$ を分節化して得られたすべての行動 $\mathbf{x}_{\beta n j}$ ($j = 1, 2, \dots, J_n$) をそれらが分類されているクラスから除き、各クラスのガウス過程のパラメータ $\mathbf{X}_c$ 、遷移確率 $P(c_\beta | c'_\beta)$ 、相互関係 $P(c_\beta | c_\alpha)$ を更新する。その後、次式に従い行動、クラス、二者の行動間の時間差をサンプリングする。

$$\mathbf{X}_{\beta n}, \mathbf{c}_{\beta n}, \mathbf{d}_{\beta n} \sim P(\mathbf{X}_\beta, \mathbf{c}_\beta, \mathbf{d}_\beta | \mathbf{S}_{\beta n}, \Theta),$$

ただし、

$$\mathbf{X}_{\beta n} = \{\mathbf{x}_{\beta n 1}, \dots, \mathbf{x}_{\beta n J_n}\}, \quad (9)$$

$$\mathbf{c}_{\beta n} = \{c_{\beta n 1}, \dots, c_{\beta n J_n}\}, \quad (10)$$

$$\mathbf{d}_{\beta n} = \{d_{\beta n 1}, \dots, d_{\beta n J_n}\}. \quad (11)$$

これにより推定された観測系列 $\mathbf{S}_{\beta n}$ の全ての分節とクラスを用いて、各クラスのガウス過程のパラメータ $\mathbf{X}_c$ 、遷移確率 $P(c_\beta | c'_\beta)$ 、相互関係 $P(c_\beta | c_\alpha)$ を更新する。この操作をすべての系列に対して行う。その後グラフィカルモデルの α と β を反転させ、観測系列 $\mathbf{S}_\alpha$ に対して同様に行動の長さ、クラス、二者の行動間の時間差をサンプリングし、パラメータを更新する。これらの操作を繰り返すことで、両方の観測系列の分節点とガウス過程のパラメータを交互に最適化し、二者間の行動の相互関係を推定することが出来る。しかし、この推定では行動の長さやクラス、その時間差のあらゆる組み合わせを計算しなければならないと計算量が膨大となるため、文献 [Uchiumi 15] と同様に Forward filtering-Backward sampling を用いて効率的に計算を行う。

2.3.2 Forward filtering-Backward sampling

本稿では行動とそのクラス、時間差を隠れ変数とみなして、Forward filtering-Backward sampling により同時にサンプリングする。Forward filtering では観測系列のあるタイムステップ t を終点とする長さが k の分節がクラス c となり、その時

相手との行動のずれが d である確率を次のように計算する.

$$A_\beta[t][k][c_\beta][d] = \mathcal{GP}(\mathbf{s}_{t-k:k}|\mathbf{X}_c)p(c_\beta|c_{\alpha_{t+d}})p(d) \\ \times \sum_{k'=1}^K \sum_{c'=0}^C \sum_{d'=-D}^D p(c_\beta|c'_\beta)\alpha[t-k][k'][c'_\beta][d'] \quad (12)$$

C はクラス数, K は行動の最大の長さ, D はずれの最大値となる. $p(d)$ はガウス分布であり, クラスの遷移確率 $p(c_\beta|c'_\beta)$ と二者間の行動の相互関係 $p(c_\beta|c_{\alpha_{t+d}})$ は, 次式のようになる.

$$p(c_\beta|c'_\beta) = \frac{N_{c'_\beta c_\beta} + \lambda}{N_{c'_\beta} + C\lambda'} \quad (13)$$

$$p(c_\beta|c_\alpha) = \frac{N_{c_\alpha c_\beta} + \gamma}{N_{c_\alpha} + C\gamma'} \quad (14)$$

ただし, $N_{c'_\beta}$ と $N_{c'_\beta c_\beta}$ は, それぞれクラスが c'_β となった行動の数と, クラスが c'_β から c_β に遷移した回数を表している. また, N_{c_α} と $N_{c_\alpha c_\beta}$ はそれぞれクラスが c_α となった α の行動の数と, クラスが c_α から c_β に遷移した回数表している. k', c'_β, d' は一つ前の $t-k$ を終点とした行動の長さ, クラス, 時間差を表しており, 式 (12) において周辺化されている.

次に Forward filtering によって求めた確率に従い, Backward sampling を行うことで観測系列の全ての分節, クラス, ずれを決定することが出来る.

2.4 インタラクシヨンのルール推定

インタラクシオンにおけるルールとは行動の関係性であると定義し, 行動間の確率に基づいてルールを抽出する. Coupled GP-HSMM によって分節・分類された α の j' 番目の行動 $c_{\alpha_{j'}}$, それと対応した β の j 番目の行動 c_{β_j} から以下の確率を考える.

1. α が行動 $c_{\alpha_{j'}}$ をした時に, β が行動 c_{β_j} をする確率 $P(c_{\beta_j}|c_{\alpha_{j'}})$
2. β が行動 c_{β_j} をした後に, β が行動 $c_{\beta_{j+1}}$ をする確率 $P(c_{\beta_{j+1}}|c_{\beta_j})$
3. α が行動 $c_{\alpha_{j'}}$ をした後に, α が行動 $c_{\alpha_{j'+1}}$ をする確率 $P(c_{\alpha_{j'+1}}|c_{\alpha_{j'}})$
4. α が行動 $c_{\alpha_{j'}}$, β が行動 c_{β_j} をした後に, α が行動 $c_{\alpha_{j'+1}}$, β が行動 $c_{\beta_{j+1}}$ をする確率 $P(c_{\alpha_{j'+1}}, c_{\beta_{j+1}}|c_{\alpha_{j'}}, c_{\beta_j})$

インタラクシオンに時間差がある場合, これらの確率を単純に計算することはできない. しかし, 提案モデルでは, その時間差も推定しているため, これらの確率を正確に推定することが可能となる. さらに, これらの確率を用いることで, インタラクシオン内に存在するルールを学習することが出来る.

3. 実験

提案手法の有効性を検証するために, 簡単なルールを組み込んだゲームを作成し, 実際にそのゲームを用いて実験を行った. 実験には kinect を用いて取得した人の右手の位置を用い, kinect から得られる 30fps の位置座標を, 4fps ヘダウンサンプリングし使用した. この値は, 計算量を削減するため, 動作の特徴が失われない範囲で経験的に決定した.

表 1: 推定された二者間の行動の関係 $P(c_\beta|c_\alpha)$

		β の行動 c_{β_j}				
		0	1	2	3	4
α の 行動 $c_{\alpha_{j'}}$	0	0.94	0.00	0.00	0.03	0.03
	1	0.05	0.92	0.01	0.01	0.02
	2	0.09	0.01	0.89	0.00	0.00
	3	0.01	0.01	0.08	0.62	0.28
	4	0.01	0.01	0.05	0.30	0.64

3.1 方向合わせゲーム

実験には簡単なルールを含む人工的に作成したインタラクシオンゲームである方向合わせゲームを用いた. ゲームの流れは以下の通りである.

1. 互いに向かい合って右腕を正面に出して待機
2. 互いに腕を三回振る
3. 左右どちらかの方向に大きく腕を振る
4. 互いに腕を振った方向が一致したら 5, 一致しなかったら 1 に戻る
5. 互いにハイタッチをして 1 に戻る

3.2 結果

図 3 がそれぞれ (a) α の右手の手先座標の時系列データ, (b) β の右手の手先座標の時系列データ, (c) α, β それぞれの右手の手先座標を重ね合わせた時系列データ, (d)coupled GP-HSMM を用いて分節化した結果, (e) α の行動と β の行動の推定された時間差, の一部のフレームを抜粋してものとなる. 結果より, 分節された各行動 ID は以下のような行動を表していると考えられる.

0. 手を前に出し一回振る行動
1. ハイタッチをする行動
2. 手を二回振る行動
3. 手を左に振る行動
4. 手を右に振る行動

これらの行動は, 今回設定したルールと比べると想定していた行動とは違っていたが, 妥当な行動が学習できたといえる. 想定では, 手を前に出す行動と三回振る行動が抽出されると考えていたが, 実際には手を前に出し一回振る行動と, 手を二回振る行動に分節化されてしまった. しかし, 手を前に出し, 三回振る動作はこのゲームにおける一連の動作のため, どこで分節されてもルールを表現することが可能であり, 妥当な結果である. また図 3 の (d) より, 両者の動作が完全に同期しているわけではないことが確認でき, さらに (c) より, 推定された両者の行動にも時間差があることが確認できる. しかし (e) をみると, 対応している行動同士が繋がるようにずれが推定されており, 相互関係を正しく推定することが可能となっている.

表 1 が相互関係の推定により得られた確率 $P(c_{\beta_j}|c_{\alpha_{j'}})$ である. この表を見ると, α が行動 0(手を前に出す)を行った時, β も行動 0(手を前に出す)を行う確率が 94%となっている. これはインタラクシオン内におけるルールであり, 共起する行動を正しく学習できていることが分かる. 図 4 が学習された遷移確率 $P(c_{\alpha_{j'+1}}, c_{\beta_{j+1}}|c_{\alpha_{j'}}, c_{\beta_j})$ により得られたインタラクシオン内における行動の遷移である. 遷移確率 $P(c_{\alpha_{j+1}}, c_{\beta_{j+1}}|c_{\alpha_j}, c_{\beta_j})$ は α の行動と β の行動から次に起こる行動を推定しており, 図を見ると α, β がともに手を前に出し振る行動を行った後, 互いに手を二回振ることが表現できている. このように, 相手の行動との相互関係を推定することで, インタラクシオン内におけるルールを教師なしで学習することが出来る.

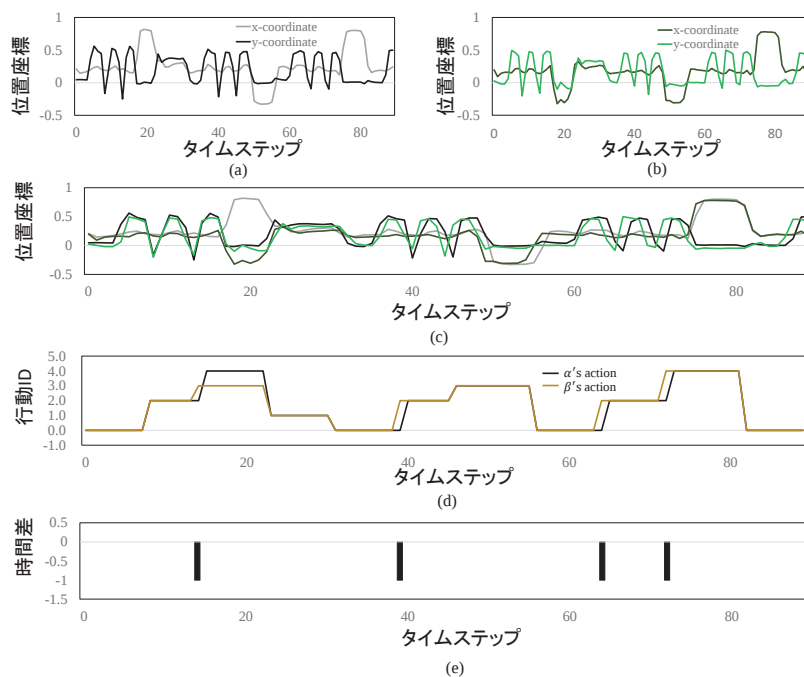


図 3: (a) α の手先座標, (b) β の手先座標, (c) α と β の手先座標, (d) α と β の分節結果, (e) 推定された時間差

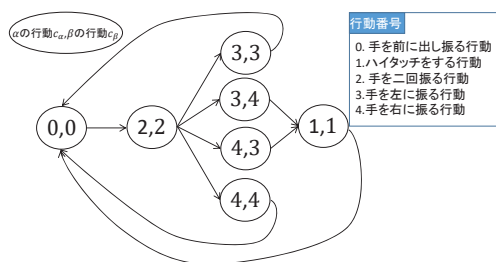


図 4: 学習されたインタラクション

4. まとめと今後の展望

本稿では、人同士のインタラクションを対象にした分節による行動の抽出とその相互関係の推定に基づくインタラクションルール学習が可能な Coupled GP-HSMM を提案し、有効性を検証するため実際のインタラクションを用いて実験を行った。実験結果より、提案モデルを用いることで二者間のインタラクションにおける動作を相互関係を推定しながら分節化することで行動を学習し、同時に行動同士のずれを推定することが可能であることが示された。また、推定された相互関係からインタラクション内のルールを確率に基づいて抽出することが出来た。今後の課題として、より複雑なインタラクションを対象にした実験が挙げられる。今回の実験は簡単な動作とルールを使用し、また行動のずれも小さかったため正しくルール抽出をすることが出来ていると考えることもできる。そのため、より複雑な動作やルールにおいても適用可能かを今後検証する。

参考文献

- [Billard 08] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, Robot Programming by Demonstration, Springer Handbook of Robotics, pp.1371-1394, 2008
- [Berenna 09] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning, A survey of robot learning from demonstration, Robotics and autonomous systems, 57(5), 469-483, 2009
- [Schaal 99] Stefan Schaal, Is imitation learning the route to humanoid robots?, Trends in cognitive sciences, 3(6), 233-242, 1999
- [Taniguchi 10] Tadahiro Taniguchi, Hiroto Nakanishi, Naoto Iwahashi, Simultaneous estimation of role and response strategy in human-robot role-reversal imitation learning, The 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, 2010
- [Soltoggio 13] Andrea Soltoggio, Felix Reinhard, Andre Lemme and Jochen Steil, Learning the Rules of a Game: Neural Conditioning in Human-Robot Interaction with Delayed Rewards, IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics, pp. 1-6, 2013
- [Matsubara 14] Y. Matsubara, Y. Sakurai, and C. Faloutsos, Autoplait: Automatic mining of co-evolving time sequences, ACM SIGMOD International Conference on Management of Data, pp.193-204, 2014.
- [Fox 11] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, Joint modeling of multiple related time series via the beta process, arXiv preprint arXiv:1111.4226, 2011.
- [Nakamura 17] Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, Hideki Asoh and Masahide Kaneko, Segmenting Continuous Motions with Hidden Semi-Markov Models and Gaussian Processes, Front. Neurorobot, vol.11, no.67, 2017
- [Uchiumi 15] Uchiumi, K., Hiroshi, T., and Mochihashi, D., Inducing Word and Part-of-Speech with Pitman-Yor Hidden Semi-Markov Models, in ACL-IJCNLP, pp. 1774-1782, 2015