

子どもたちの主体性や成長の可能性を引き出すヒューマンロボット インタラクションの実現に向けて

Toward Realization of Human-Robot Interaction that Elicits Children's Independence and Potential for Growth

岡 夏樹[†], 塚本 亜美[†], 寺岡 弘貴[†], 鶴田 穰士[†], 廣田 敦士[†], 早川 博章[†]
市川 淳[†], 田中 一品[†], 長井 隆行[‡], 大森 隆司[※]

Natsuki Oka, Ami Tsukamoto, Koki Teraoka, Joji Tsuruda, Atsushi Hirota, Hirofumi Hayakawa

Jun Ichikawa, Kazuaki Tanaka, Takayuki Nagai, Takashi Omori

[†] 京都工芸繊維大学, [‡] 電気通信大学, [※] 玉川大学

Kyoto Institute of Technology, The University of Electro-Communications, Tamagawa University

nat[atmark]kit.ac.jp

Abstract

This paper describes a plan to create a robot that emergently interacts with children using a recurrent neural network and draws out children's identity and potential for growth. Preliminary experimental results of basic functions of emergent interaction are also reported.

Keywords — Emergent interaction, Recurrent neural network, Intrinsic motivation

1. はじめに

我々は、保育園での身体表現活動の現場で収集したデータを分析し[1]、子どもたちの社会的インタラクションの発達をモデル化する試みを継続中であり、将来、保育の支援に活かすことを目指している。この身体表現活動の指導者である潮木は、文献[2]の中で次のように述べている：「身体表現の活動は、身体の動きのポキャブラリーを発達させることにより、子どもたちの生まれつきの能力を引き出し、可能性を高め、自信を持たせることが目的である。指導場面では、自然と表現を楽しむ子どもたちがいる一方で、自己の感情を表現することへの戸惑いを感じ、活動に意欲的でない子どもたちにも出会ってきた。その中で、身体で表現することを子どもたち自身が楽しむためにはどのような援助・指導の方法があるのかを探る日々であった。」

「自らの主体的な表現を、肯定的に捉えてもらえる環境の中で、それぞれの表現方法に沿った援助が行われ、子どもたちが人とのかかわりあいを楽しむこと、子どもたちがモノとのかかわりあいを楽しむことが、子どもたちがこととのかかわりあいを楽しむ経験が、表現を楽しむ力の基礎になると考えた。」

本研究の目的は、子どもたちの仲間として身体表現活動に参加するロボットを開発し、身体表現活動への参加に意欲的でない子どもが参加するきっかけや、身

体表現を楽しむきっかけを提供したり、身体表現活動の場がより創造的になるよう支援したりする可能性を探ることである。具体的には、ロボットが受動的な動きと能動的な動きを適切に切り替えることにより、創造的な場の形成に資することを目指している。受動的な動きとして、子どもの主体的な表現をロボットがまねることにより、その表現を肯定的に捉えていることを示したり、能動的な動きとして、子どもたちの定型動的な動きが続いたとき、新しい動きを主導したりすることを計画している。

その第一歩として、指導者による太鼓やピアノのリズムに合わせて歩行風に動く家庭用掃除ロボット(図1)の制作に着手した[3]。文献[3]のロボットは、あらかじめプログラムされた2通りのモード(人が出す音に同期してロボットが動くSynchronyモードと、各種の音楽におけるCall and Responseのように人とロボットが交互に掛合いをするTurn-takingモード)のうちの指定したモードでリズムに合わせて動くものであった。



図1：ロボットの動作

これに対して、本論文では、身体表現活動の場へのロボットの投入により、子どもたちのより豊かな成長の機会が提供できる可能性を探ることを目指し、2節にて、より創発的なインタラクションを可能にするロボ

ットの実現に向けての構想を示す。続いて3節で、リカレントニューラルネットワークを用いた創発的インタラクションの基本機能の試作（計算機上での試作のみ；ロボットへの実装はまだ）について述べ、4節でその評価実験について報告する。

2. 創発的なモード遷移に向けて

本研究では、Synchrony、Turn-taking 各モードでのインタラクションの仕方を設計してあらかじめロボットに与えておくのではなく、学習により典型的なインタラクションの仕方（各モードでの典型的な行動）を獲得させておき、また、適切に設計した内部報酬を定めておくことで、人との相互作用の中で創発的にインタラクション・モードが立ち現れることを目指す。人とロボットのどちらが主導的にインタラクションするかの移り変わりも創発的に生じることを目指す。

2.1 リカレントニューラルネットワークによるモデル化

文献[3]の実装ではインタラクション・モードをキーボードからの入力により明示的に切り替えたが、本研究では下記で紹介する研究[4]を参考にして、Recurrent Neural Network with Parametric Biases (RNNPB) を用いて創発的なモード遷移を実現することを目指す。

Ito と Tani は、人型ロボットと人が、互いに動きを模倣しあうインタラクションを RNNPB を用いて実現した[4]。まずロボットに、人の周期的な動きとそれに対応するロボットの周期的な動きの時系列パターンを複数学習させた。複数のパターンは、それぞれ異なる Parametric Biases (PB) の値と対応する時系列パターンとして学習された。学習後、人とのインタラクション実験を実施し、学習済みの人の動作パターンを与えたときのロボットの動きや、学習時とは異なる人の側の新しい動きに対するロボットの動きが調べられた。学習済みの人の動作パターンを模倣できたこと、および、人の新しい動作パターンに対しては、1) 学習済みの類似パターンの動きを生成する場合、2) 学習パターンとは異なる新しい周期的な動作パターン（人の動きと同期）を生成する場合、3) 不安定な動きをする場合—があることが報告された。

本研究で対象とするヒューマンロボットインタラクションを RNNPB で実装する場合の模式図を図2に示す。図中の r はロボットの動き、 h は人が出す音に、

それぞれ対応するノード、 p は PB ノード、 c はコンテキスト・ノードであり、添え字の t 、 $t+1$ は時刻を示す。学習時は Synchrony、Turn-taking 各モードにおける r と h の時系列の例が与えられ、それらを予測できるようにリンクの重みと PB 値の学習が進む。インタラクション中は観測される h_{t+1} との誤差に基づき PB 値が変化し、それに応じて r_{t+1} が決まる。

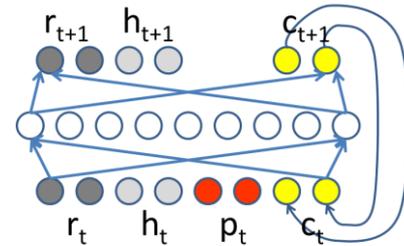


図2: RNNPBの模式図

2.2 内発的動機づけの導入計画と予想される効果

我々は、以上のような感覚運動系の予測学習に基づくモード遷移に加えて、ロボットに内発的動機づけ[5、6、7、8]を導入して、ロボットと人の間の創発的なインタラクションを実現することを検討している。ロボットは「外部報酬と（内発的動機づけに対応する）内部報酬の和」の期待値が大きい行動を選択する。本研究では内部報酬として、1) Synchronyの心地良さに対応する報酬、2) Turn-takingの心地良さに対応する報酬、3) 真似できたうれしさに対応する報酬、4) 真似されたうれしさに対応する報酬、5) 新奇性・親近性への選好に対応する報酬（予測誤差の変化に基づき計算；後述）等を導入することを予定している。これらの導入により、インタラクション・モード間の創発的な遷移や、インタラクションの主導権の自発的な移り変わりが実現できると考えている。

今後のシステムの実装は、モジュール組換を基本演算とするアーキテクチャ[9]（図3）を使用する計画である。ここでのモジュールのいくつかは、時系列の予測・学習機能を持つものであり、PBを持たない通常のリカレントニューラルネットワーク(RNN)またはRNNPBで実装する予定である。予測誤差の変化（図4）に応じて、内部報酬が多くなるように、当該予測モジュール、および、関連するモジュールのゲートの開閉（モジュールの切替）を、以下の通り行うことで、新奇性・親近性への選好を表現する計画である。

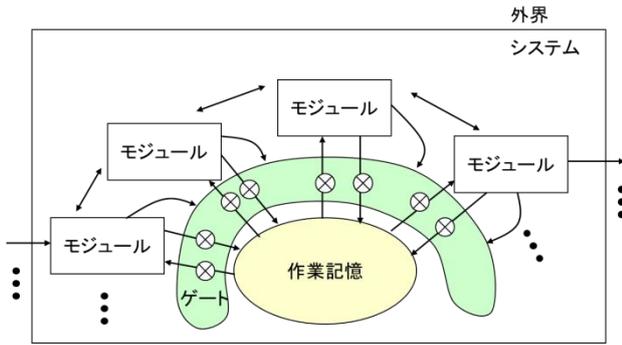


図3：モジュール組換アーキテクチャ

- A. 予測誤差が大きいままのモジュール：内部報酬が生じず、結果として内部報酬を生み出す他のモジュールに切り換わる。学習できそうにないので飽きた状態に相当。
- B. 予測誤差が増加中のモジュール：内部報酬が生じず、結果として内部報酬を生み出す他のモジュールに切り換わる。予測対象（相手も含む）が変化したと考えられるため、その変化に応じた行動や内部情報処理をするモジュールに切り換わる。
- C. 予測誤差が減少中のモジュール：内部報酬を生み

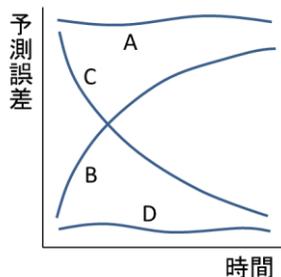


図4：予測誤差の変化のパターン（4通り）

出し続けるため、そのままそのモジュールを選択し続ける。学習進行中であるので、それへの選好（親近性への選好）を保ったまま学習を継続することに相当。

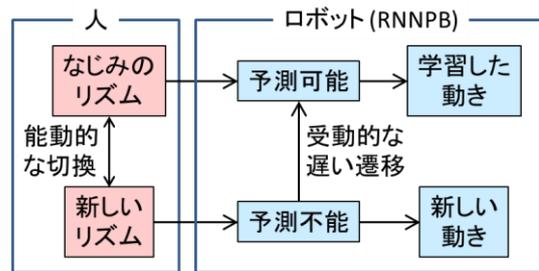
- D. 予測誤差が小さいままのモジュール：内部報酬が生じず、結果として内部報酬を生み出す他のモジュールに切り換わる。学習完了しているため、飽きて別の行動や内部情報処理をすること（新奇性への選好・好奇心）に相当。

なお、RNNPB をモジュールとして用いる場合は、モジュール内に複数の時系列パターンが分散的に埋め込まれているため、これらの選択に内部報酬による選好を加味するためには、その選好を何らかの方法で予

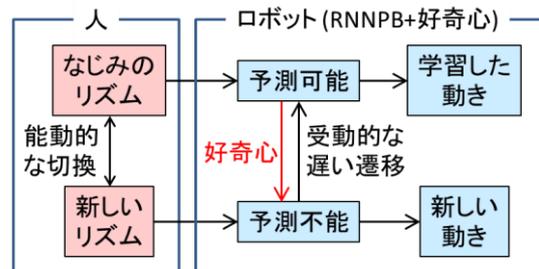
測誤差に換算して PB の値を介して制御することになるだろう。

RNNやRNNPBによる対人インタラクションでは、基本的には人の側が主体的な切替をして（図5(a)左）、それをきっかけにロボットが対応した動作をすることになる（図5(a)右）。このため、概ね、人主導のインタラクションが起こると思われる。ただし、人の動きが学習済みのものでなかった場合は、ロボットが新しい動きを始める（図5(a)右下）可能性があり、それがロボット主導の動きのように感じられる可能性がある。また、ロボット主導のインタラクションが人の側の動きをきっかけに開始されるような時系列をあらかじめ学習させておくことは可能であると考えられる。ただし、これら2つのケースでは、ロボット主導に感じられる動作が行われた場合も、そのきっかけを与えたのは人であることに注意すべきである。

これに対して、内発的動機づけを付加したモデルでは、内部報酬をきっかけとしてロボットが動作する（図5(b)）ことにより、（見かけ上でなく）真にロボット主導でインタラクションを開始できると考えている。具体的には、図4のDのようなケースでは、誤差を小さくするという規準で動作する学習システムは、現状を維持し続けるが、好奇心や飽きに相当する内発的動機を備えたシステムは、主体的にモジュールを切り換えることができるが、その結果予測不能に陥り、新しい動きを表出することになる。



(a)



(b)

図5：好奇心の追加の効果

リトミックにおいては、子どもが主体的な表現をした場合は、ロボットは少しの間の不安定な遷移を経て、それに追従する動きをすると予想している。これにより、子どもは自らの動きを肯定的に捉えてもらったと感じることを期待する。一方、子どもたちが受動的で定型的な表現を続けた場合は、ロボットは好奇心に基づき、強制的にモジュールを切り換え、結果として新しい動きをして、リトミックの場に刺激を与えることを期待する。以上のように、ロボットは自らが置かれた場の性質に応じた動き（主体的な動きに対してはそれを支持し、逆に受け身の動きに対しては自らが新たな刺激を生み出す）をすることにより、より豊かで子どもの成長を引き出すヒューマン・ロボット・インタラクションを実現できる可能性が高まると思われる。

3. RNNPB を用いた創発的インタラクションの基本機能の試作

RNNPB は通常の RNN と同様に BPTT (Back Propagation Through Time) 学習法を用いるが、ネットワークの重みを更新するタイミングで PB 時系列の更新も行う[4]。PB 時系列とは学習する時系列の各ステップに対する PB 値の列であり、BPTT で PB ノードに伝搬した入力誤差を用いて次の更新式(1)(2)(3)によって計算される。

$$\delta\rho_t = k_{bp} \cdot \sum_{step=t-\frac{d}{2}}^{t+\frac{d}{2}} \delta_{step}^{bp} + k_{nb}(\rho_{t+1} - 2\rho_t + \rho_{t-1}) \quad (1)$$

$$\Delta\rho_t(s) = \varepsilon \cdot \delta\rho_t + \eta \cdot \Delta\rho_t(s-1) \quad (2)$$

$$p_t = \text{sigmoid}(\rho_t) \quad (3)$$

ここに、 ρ_t はステップ t の PB ノードの内部値、 p_t はステップ t の PB 値である。

式(1)の第1項は各ステップの PB ノードに伝搬した誤差信号 δ^{bp} を一定ステップ分だけ足し合わせたものであり、局所的な誤差の影響ではなく系列が構造的に変化したときにだけ PB 値が変化ようになる。式(1)の第2項は前後ステップの値を用いて平均化を行い、PB 値が急激に変化するのを防いでいる。

式(2)は式(1)で求めた $\delta\rho_t$ に基づいて最急降下法を用いて学習 $s-1$ ステップから s ステップへの PB 時系列の内部値を更新する仕方を示している。式(2)で、 ε は学習率、 η は慣性項の係数である。

学習において全ての教示パターンについて予測エラ

ーが最小化された場合、PB 時系列から教示パターンへのマッピングが自己組織化される。よって、収束した PB 値のみを RNN に入力する(入力信号は初期値のみを与え、次のステップからは自らの予測信号をフィードバックさせる)と PB 値に対応したパターンを生成することができる。

ここまでは、全時系列の入力と教師信号がわかっているオフラインでの RNNPB の学習について述べてきたが、本研究では PB ノードへの入力エラーから PB 値を更新する計算が 1 ステップごとにリアルタイムにできる必要がある。オフラインでは計算したいステップから前後数ステップ分のデータを用いて PB の計算を行っているが、オンラインでは過去のステップのデータのみで計算を行わなければならない。また、インタラクションの性質上計算時間に制約が生じるため、過去の全てのステップのデータを用いて計算するわけにもいかない。そこで、大谷[10]の方法に倣い、参照窓範囲を設定し、その範囲内の各ステップに対してオフラインでの PB の計算手法を適用し、参照窓範囲内のステップ分の PB 値を平均化したものを次ステップでの PB 値とするリアルタイム計算手法を用いた。まず W ステップ前から現在までの出力誤差列を逆伝播させて PB ノードへの入力エラーの列を求める。PB ノードへの入力エラーの列と W ステップ前から現在までの PB 時系列から、前述の式(1)~(3)を用いて学習 1 ステップ分の PB 時系列の更新を行う。そして、更新された W ステップ前から現在までの PB 時系列の平均をとってステップの PB 値とする。過去の時系列から求めた PB 値とリアルタイムに入力された入力信号から次のステップの予測を行い、次のステップで入力信号が入ってきたら予測と比較して出力誤差を計算する。出力誤差が求めたら、参照窓範囲をずらして、 W ステップ前から現在までの出力誤差列から PB ノードへの入力エラー列を求め、参照窓範囲の PB 時系列更新を行う。このように、参照窓範囲をずらしながら更新した PB 時系列の平均値を次のステップの PB 値にすることを繰り返す。

4. 創発的インタラクションの基本機能の評価実験

本節では、通常の RNN と比較しながら、RNNPB の訓練例に対する学習結果と、未学習パターンへの反応を調べる。

4.1 実験設定

本実験における Synchrony モードは、人が演奏した 1 小節のリズムパターンから次の小節のリズムパターンを予測することにより、音と同期して動くものとする。Turn-taking モードは、人が演奏した 1 小節のリズムパターンと同じパターンで次の小節の間動く（人はその間休止する）ことを繰り返すものとする。

1 小節のリズムパターンは、8 分音符を最小単位とした 1/0 の記号列により表現する。4 分の 4 拍子の 1 小節内の 8 つの時点で、音符がある箇所は 1、そうでない箇所は 0 と表記される。図 6 にこの方式によるリズムパターン表現の例を示す。



図 6：リズムパターンの表現方法（例）

ニューラルネットワークの構成（図 2 を参照）は、次の通りである：

RNNPB: h ノード 8, r ノード 8, p ノード 2, c ノード 18, 中間層の隠れノード 30

また、PB の更新式(1)(2)(3)のパラメータの値は次の通り：

$$k_{bp} = 0.2, \quad k_{nb} = 0.1, \quad d = 10$$

$$\varepsilon = 0.1, \quad \eta = 0.9$$

比較した RNN は、上記の構成から PB を除いたものである。

学習データは、「さくらさくら」「もみじ」「きらきら星」「ちょうちょう」「ドナドナ」の 5 曲のリズムパターンを、Synchrony、Turn-taking の両モードに対して用意した。

Synchrony モードの学習データでは、入力層の h ノード（人の演奏に対応）と r ノード（ロボットの動きに対応）への入力共通で、楽譜通りに 1 小節ずつ、順次入力される。出力層の h ノードと r ノードからの望ましい出力（教師信号）も共通で、1 小節先のリズムパターンである。

Turn-taking モードの学習データでは、入力層の h ノードと r ノードへの入力は交互にすべて 0 となる。最初のステップでは、h ノードには 1 小節目のパターンが、r ノードにはすべて 0 のパターンが入力され（すなわち、人が 1 小節演奏する間、ロボットは待機）、第 2 ステップでは、h ノードにはすべて 0 が入力され、r ノードには 1 小節目のパターンが入力される（すなわち、ロボットは前ステップの人の演奏を模倣し、人は

その間待機する）。第 3 ステップでは、h ノードには 2 小節目のパターンが、r ノードにはすべて 0 のパターンが入力され、以下同様に続く。出力層の h ノードと r ノードからの望ましい出力（教師信号）は、次のステップのそれぞれの入力パターンであり、交互にすべて 0 となる。

以上の、5 曲×2 モード=10 種類の訓練データの順番をシャッフルし、1 番目のデータをニューラルネットに入力して、ニューラルネットの重みと PB 値を更新、続いて 2 番目のデータを……。これを 10 番目のデータまで行うことを 1 エピソードとし、10000 エピソード繰り返して学習した。

また、未学習データとして、学習データに含まれない「春が来た」「富士山」の 2 曲のリズムパターンを、Synchrony、Turn-taking の両モードに対して用意した。

4.2 訓練例に対する学習結果

10 種類のどの訓練データに対しても、概ね正しい出力を予測できるようになったが、RNNPB の予測精度は、RNN と比べて若干落ちる傾向があることが分かった。

10 種類の訓練データに対する PB 値の分化の様子を図 7 に示す。似た時系列パターンの曲（「さくら」と「ちょうちょう」）の PB 値が近いことや、同じ曲の Synchrony、Turn-taking の両モードに対する PB 値が比較的近い位置にあることや、Synchrony モードの 5 曲の分布と、Turn-taking モードの 5 曲の分布が、それぞれある程度まとまりがあることが分かる。

また、PB 値から学習済みパターンを再生できるかを調べるため、PB 層に学習で得られた PB 値を入力し、入力層には初期値のみを与えて次の時刻からは自らの予測をフィードバックさせる閉ループによる計算を行った。その結果、10 種類のデータのうち、8 種類については、学習した時系列パターンが生成できていることが分かった。学習パターンが生成できなかった 2 種類のデータ（「さくらさくら」Synchrony モードと、「ちょうちょう」Synchrony モード）を調べると、この 2 つが汎化された同じパターンが生成されていることが分かった。この 2 種類のパターンに対する PB 値が非常に近い値であることが図 7 から分かる。

4.3 未学習パターンへの反応

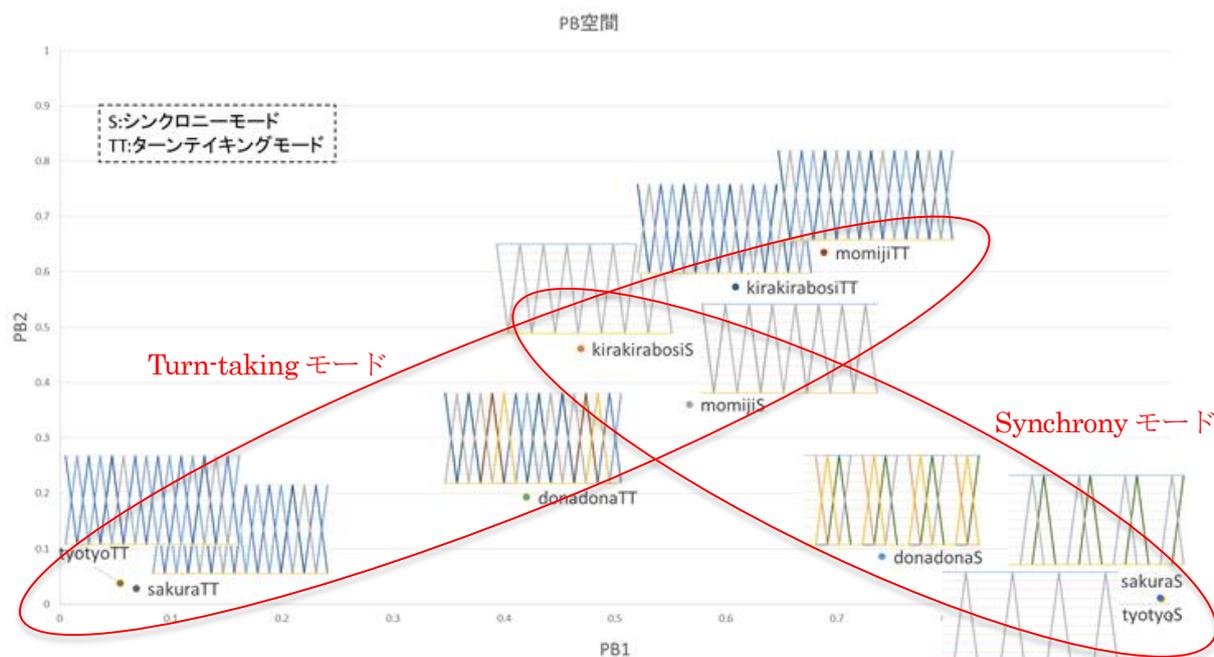


図7：10種類の訓練データに対するPB値の分化の様子

未学習データである4種類のリズムパターン（「春が来た」Turn-takingモード、「富士山」Turn-takingモード、「春が来た」Synchronyモード、「富士山」Synchronyモード）から構成されるテストデータ（この記載順に各パターン50ステップずつの人側のリズムデータから構成した）を作成し、この入力に対するロボット側の出力動作を調べた。

RNNでは適切なモード切り替えが起こらなかったのに対して、RNNPBでは、人の側がTurn-takeからSynchronyに切り換わった辺りで、両モードが混ざったような興味深い状態を経由して、Synchronyモードに切り替わる様子が観察された。両モードが混ざったような状態としては、次の2種類が観察された：1) 待機状態と動作状態が混ざったような0が7個で1が1個だけのパターン、2) 人もロボットも同時に動作する点ではSynchronyだが、ロボットの動作が1小節前の人のまねになっている。

この実験時のPB値の変化のグラフ（図8）を見ると、最初はPB1とPB2の値が近いというTurn-takingモードに共通して見られる特徴（図7）を示していたが、人の側のモード切り替えから少し遅れてPB1とPB2が異なる値となり、その変化の過程で上述の両モードが混ざったような動作を示していたことが分かる。

以上のことから、RNNPBの学習は、RNNの学習と比べて、2種類のモードからなる訓練データの特徴をうまく捉えたものとなっていたため、未知のリズムパ

ターンに対しても適切にモード切り替えができ、また、モード切り替えの過程で興味深いふるまいが創発したと考えられる。

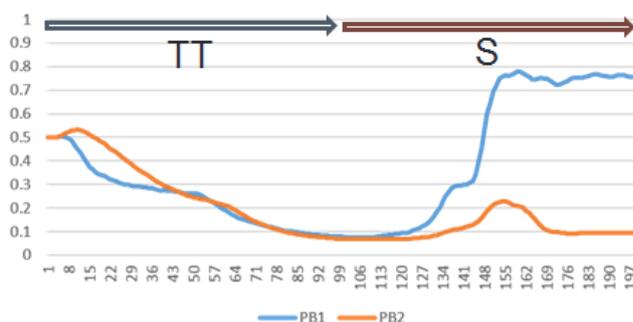


図8：未知パターンに対するPB値の変化の様子

5. まとめ

本論文では、RNNPBを用いて、未知パターンに対する創発的な反応が観察できたことを報告した。これにさらに内発的動機づけを付加し、ロボットに実装することにより、身体表現活動の場の様子に応じた適切な動作（主体的な動きに対してはそれを支持し、逆に受け身の動きに対しては自らが新たな刺激を生み出す）を可能にしたい。

このような適応的な動作を可能にしたロボットを身体表現活動の場に導入し、子どもたちの行動、インタラクション、社会的な関係性がどのように変化するかを、1) ロボット無し、2) 受動的にまねをするだけの

ロボット、3) 常に主体的に(勝手に)ふるまうロボットと比較して、観察・分析する計画である。

なお、今想定している掃除ロボットの歩行様の動作は、現在協力先の保育園で実施中のリトミックにおける身体表現活動の一部(リズムに合わせて歩いたり走ったり止まったりする)だけに対応したものである。止まった時に様々なポーズをとる活動、ボール、リボン、布など、モノを使った活動への対応は今後の課題である。

謝辞

本研究は科研費(25330260 および 26118003)の助成を受けたものである。

参考文献

- [1] Zhang, B., Nakamura, T., Ushioji, R., Nagai, T., Abe, K., Omori, T., Oka, N. & Kaneko, M., (2016) "Simultaneous children recognition and tracking for childcare assisting system by using Kinect sensors", *Journal of Signal and Information Processing*, Vol. 7, pp. 148-159.
- [2] 潮木玲奈, (2016) "幼児期における身体を通じたかかわりあいの発現", 玉川大学大学院教育学研究科修士論文.
- [3] 寺岡弘貴, 高橋卓見, 石崎祐太, 高橋ともみ, 塚本亜美, 西村宏武, 早川博章, 廣田敦士, 古橋翔吾, 松尾星吾, 岡夏樹, (2016) "踊 Roomba ~インタラクション・モードの創発~", 2016 年度情報処理学会関西支部 支部大会, B-102, 3 pages.
- [4] Ito, M. & Tani, J., (2004) "On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system", *Adaptive Behavior*, Vol.12, No.2, pp.93-115.
- [5] Schmidhuber, J., (2010) "Formal theory of creativity, fun, and intrinsic motivation (1990-2010)", *IEEE Transactions on Autonomous Mental Development*, Vol.2, No.3, pp. 230-247.
- [6] Gottlieb, J., Oudeyer, P.Y., Lopes, M. & Baranes, A., (2013) "Information-seeking, curiosity, and attention: Computational and neural mechanisms", *Trends in Cognitive Sciences*, Vol.17, No.11, pp. 585-593.
- [7] Kidd, C. & Hayden, B.Y., (2015) "The psychology and neuroscience of curiosity", *Neuron*, Vol.88, No.3, pp. 449-460.
- [8] 高井利将, 岡夏樹, 早川博章, (2016) "インタラクションを通じた数の概念の獲得", 2016 年度 人工知能学会全国大会(第30回) 論文集, 104-OS-22a-3, 4 pages.
- [9] Oka, N., (2002) "Apparent "free will" caused by representation of module control", *No Matter, Never Mind. Proceedings of Toward a Science of Consciousness: Fundamental Approaches*, pp. 243-249.
- [10] 大谷拓, (2005) "人とロボットの協調作業のためのRNNPBによる擬似シンボルの獲得", 京都大学工学部情報学科特別研究報告書.