

ロボットとの協調作業のための RNN による擬似シンボルの獲得

大谷拓[†] 大庭隼人[†] 駒谷和範[‡] 尾形哲也[‡] 谷淳^{*} 奥乃博[‡]

[†] 京都大学工学部情報学科 [‡] 京都大学大学院情報学研究科知能情報学専攻 ^{*} 理化学研究所脳科学総合研究センター

1. はじめに

近年、多くのヒューマノイドロボットが開発され、人間とロボットが協調作業をとることが必要になってきた。人間と機械が協調して作業を行なうためには、両者間における情報の共有が必要不可欠となる。共有される情報を大きく分けると、動作レベルでの情報と言語レベルでの情報の2つに分けることができる。動作レベルでの情報とは、現在の運動情報を指し、作業に身体的動作を含めば含むほどその重要性は高くなる。しかし、運動情報は常に動的に変化するので、その情報をお互いが把握することは難しい。一方、言語レベルでの情報とは、人間同士の対話協調で用いるシンボリックな情報を指し、その抽象度の高さによって非常に高度なインタラクションを可能にする。しかし、タスクを明示的に言語で表現すること自体が難しい。従って協調作業のためには、動作ほど具体的ではなく言語ほど抽象度の高くない、新たな共同注意対象が必要となる。

本研究では近年盛んなロボットによるタスク表現(擬似シンボル)の研究に着目し、ロボットに学習により擬似シンボルを獲得させ、それを動作生成のみならず人間とのインタフェースに応用する方法について検討する。このようなインタフェースが実現すると、コミュニケーションの円滑化および作業の効率化が期待できる。これまでに尾形ら [2] は、ナビゲーションタスクを設定した実験によりいくつかの擬似シンボルの獲得および、擬似シンボルを使ったインタラクションを行なっている。本稿では、谷 [1] により提唱された RNNPB (Recurrent Neural Net with Parametric Bias) の手法をロボットに適用し、人間とのコミュニケーションに RNNPB のパラメータ表現(動作プリミティブ)を利用する方法について報告する。設定したタスクは、ロボットと人間による卓上物体の協調作業であり、各方向への移動、回転などの動作を RNNPB のパラメータ表現による擬似シンボルとして獲得を試みた。

2. 動作プリミティブと RNNPB

近年ミラーニューロン発見を契機に動作プリミティブ(運動の構成単位)に基づいた運動認知、生成に関する研究が盛んになりつつある。最も代表的な手法として川人らの MOSAIC [3] がある。これは複数の予測学習器が、異なった動作プリミティブを保持し(“エキスパート”)、各々の予測信頼度に基づいて現在の環境認識・動作生成の制御を分配する手法である。各予測器が独立に表現されている点の特徴である(“局所表現”)。これに対し谷は、一つの予測学習器が複数の動作プリミティブを保持し、その切替をパラメータにより行う Jordan 型の RNN にパラメータバイアス(以下 PB)を追加したもの(RNNPB)を提案している(“分散表現”)[1]。記憶容量は小さくなるが、複数のプリミティブ間の関係性が一つの学習器の中に構成され、結果として新しい動作プリミティブを創発する

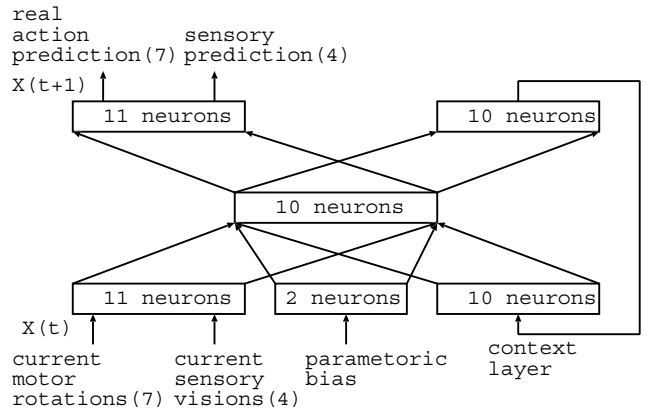


図 1: RNNPB

能力を有する。

ロボットに実装した RNNPB(図 1)は、現在のセンサ・モータ状態を入力とし、次のセンサ・モータ状態を出力(予測)する(入力・出力共に 11 ニューロン)。コンテキスト層には 10 ニューロン、PB 層には 2 ニューロンを用意した。

2.1 BPTT(Back Propagation Through Time) 学習法

BPTT(Back Propagation Through Time) 学習法により得られる PB 入力エラーを時間的に移動平均して PB 値の修正量を算出する。この結果、PB 値の変動(0~1 の連続値)は入力層、出力層のダイナミクスを“緩やか”に表現したものになる。PB 値変化の過渡期を短くすることで、長期シーケンスを「RNN の予測能力に基づいて分節化した表現」が得られる。

2.2 PB 値の生成式

$$\delta \rho_t^i = k_{bp} \sum_{step=t-d/2}^{t+d/2} \delta_t^{bpi} + k_{nb} (\rho_{t+1}^i - 2\rho_t^i + \rho_{t-1}^i) \quad (1)$$

$$p_t^i = \text{sigmoid}(\rho_t) \quad (2)$$

この 2 つの式によって PB 値が計算される。(1) 式の第 1 項の δ_t^{bpi} はステップ t における出力の誤差をバックプロパゲーションにより i 番目のパラメータノードに逆伝播したものであり、その値を固定区間 d の間で足し合わせることによって、局所的な誤差の影響を緩和させている。これによって、動作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。(1) 式の第 2 項は PB 値自体が急激に変化するのを防ぐために前後ステップの値を用いて平均化を行なっている。そして (2) 式で、シグモイド関数を通して PB 値を生成する。

3. 擬似シンボル獲得実験

3.1 擬似シンボル

本研究での擬似シンボルとは、前記の PB 値を閾値(0.5)処理によってシンボル(0, 1)に落したものを指し、この擬似シンボルを介して人間とロボットがインタラク

ションを行う。具体的には、ロボットが学習によって得られたの PB 値を音声 (手で与えた動作番号) に変換し外部出力し、それにより協調者がロボットの内部状態の理解および擬似シンボルの意味付けを行いインタラクションを試みる。ここで PB 値を閾値処理した理由は、PB 値を連続値としてそのまま出力すると人間側の理解が困難になるという点にある。また閾値処理により情報量を落している点から“擬似”シンボルと本稿では呼んでいる。

3.2 協調タスク

本研究のテストベッドは ATR 製のロボビーを改良したものであり、頭部にモーター 3 自由度、腕部にモーター 4 自由度がある。センサーは、頭部にマイクとカメラを搭載している。実験を行ったタスクは図 3 の卓上物体を人間とロボットがお互いの片手で押したり、回転させたりして移動させていくというものを設定する。その際、カメラからの画像情報を元に物体の赤色、青色部の中心座標を抽出し、カメラの中心に物体の中心が一致するように頭部のモータを動かす。

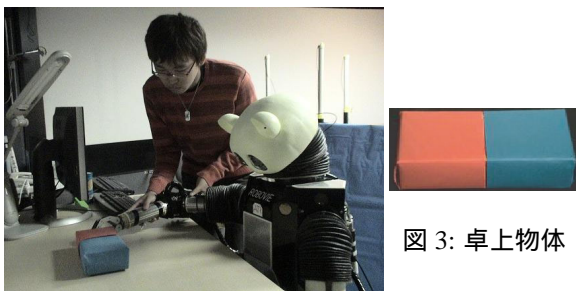


図 2: 教示風景

図 3: 卓上物体

3.3 分節化実験

図 2 のように直接ロボットに人間が手で動作を教示することにより、タスク達成に必要な学習データを収集させる。学習データとしては、モータ情報である右腕 4 関節の角度、頭部 3 関節の角度、および卓上物体の位置情報である赤色、青色それぞれのカメラ画像内 2 次元中心座標、総計 11 次元のデータを各ステップ (0.1 秒) 毎にサンプリングする。今回は協調作業を行うための実験として、以下の 1 つのパターンの実験を行った。1 パターンのステップ数は 50 (5 秒) であり、このステップ内に一度だけ各動作を行っている。また 5 秒間の動作の中には静止動作と物体初期位置への首振り動作が含まれている。

- 動作 1: 物体を右から左に押す (その後静止)
- 動作 2: 物体を手前から奥に押す (最初に首振り)
- 動作 3: 物体を右から左に押し、手前から奥に押す

4. 結果と考察

3 パターンの学習データ (各 50 ステップ) を 1 つの RNNPB を用いて反復学習を行なった。学習回数、3 パターン (150 ステップ) 分の BPTT 学習を 1 回とし、合計 10 万回の学習を行なった。図 4 のグラフは、学習後の RNNPB に動作 3 のデータを入力した場合の各ステップ毎のデータ値および PB 値を表すものである。ここではグラフを見やすくするために、入力層の 11 個の入力値のうち代表的な 7 個だけを載せている。

また表 1 は、パラメータノードを 2 つとした場合の学習によって得られた PB 値と動作との対応表である。こ

れは動作 1、および動作 2 のグラフの PB 値と動作 3 のグラフの PB 値の対応を取ったものであり、パラメータノード 2 つにより 4 つの動作に分節化できたことが分かる。つまり今回の実験により、「右から左に押す」「手前から奥に押す」「静止する」「物体の初期位置へと首を振る」という 4 つの動作を擬似シンボルとして獲得できたと言える。またパラメータノードの数を 3 つに増やした場合には、動作 3 における「右から左に押す」動作と「手前から奥に押す」動作の 2 動作間のやや複雑な動作である「遷移動作」が新たな動作として分節化され、4 つにした場合にはその新たな「遷移動作」がさらに 2 つの動作に分節化されることが確認された。

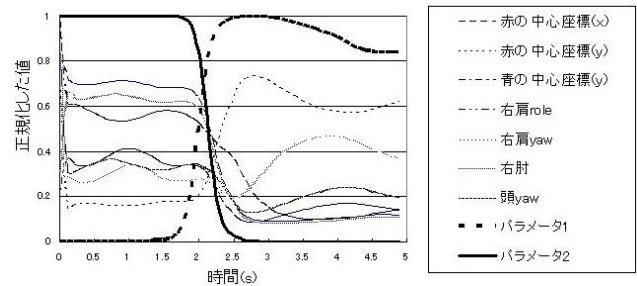


図 4: 学習後の動作 3 の出力系列および PB 値

表 1: PB 値と動作の対応

PB1, PB2	動作
0, 1	右 左 (動作 1)
1, 0	手前 奥 (動作 2)
1, 1	静止動作
0, 0	初期位置への首振り動作

5. おわりに

本研究により協調作業を行うために必要な動作を、RN-NPB を用いて擬似シンボルとして獲得できることが確認された。学習データからの擬似シンボル獲得数は、パラメータノード数だけでなく、PB 値の計算時に用いる固定区間 d および設定する閾値によっても変化するので、さらに検討を続ける必要がある。また卓上物体を使っただけの協調作業を考えた場合、他にも「右回転」「左回転」「回転しながら奥に押す」などといった動作が考えられ、今回の実験で得られた動作パターンだけでは不十分であると言える。今後さらに学習データを収集し、新たな動作パターンを獲得させる必要がある。また今回の実験では擬似シンボルの獲得を確認しただけであるので、今後は人間との協調作業における擬似シンボルの有効性について評価する必要がある。

謝辞 本研究の一部は、科研費、21 世紀 COE、大川情報通信基金、理研の支援を受けた。

参考文献

- [1] Jun Tani, Masato Ito: "Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics: A Robot Experiment", *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, Vol.33, No.4, pp.481-488, 2003.
- [2] Tetsuya Ogata, Masaki Matsunaga, Shigeki Sugano, and Jun Tani: "Human Robot Collaboration Using Behavioral Neural Networks", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1592-1597, 2004.
- [3] M.Haruno, D.Wolpert, and M.Kawato: "MOSAIC model for sensorimotor learning and control", *Neural Computation* 13, pp.2201-2220, 2001.