The University of Osaka
Institutional Knowledge Archive

Title	予測符号化モデルによるシェマの発達過程の解明
Author(s)	
Citation	令和5（2023）年度学部学生による自主研究奨励事業研究成果報告書. 2024
Version Type	VoR
URL	https://hdl.handle.net/11094/95193
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

様式６

申請先学部　　基礎工学部　採択番号　No.7

令和５年度大阪大学未来基金「学部学生による自主研究奨励事業」研究成果報告書

ふりがな 氏　　名	のむら　けんたろう 野村　健太郎	学部 学科	基礎工学部 システム科学科	学年	3年
ふりがな 共　同 研究者氏名		学部 学科		学年	年
					年
					年
アドバイザー教員 氏名	堀井　隆斗	所属	基礎工学研究科		
研究課題名	予測符号化モデルによるシェマの発達過程の解明				
研究成果の概要	研究目的、研究計画、研究方法、研究経過、研究成果等について記述すること。必要に応じて用紙を追加してもよい。（先行する研究を引用する場合は、「阪大生のためのアカデミックライティング入門」に従い、盗作剽窃にならないように引用部分を明示し文末に参考文献リストをつけること。）				

1.　研究背景

本研究は，乳幼児の認知発達メカニズムを解明するための計算モデルの構築とそれを用いた人の認知構造の発達的変化の分析を目的とする．目的達成のため，本研究では発達心理学において重要視されている Schema 理論と脳の基本原理として近年提唱されている予測符号化に注目し，認知発達モデルを構築した．本章では本研究の背景にある Schema 理論及び予測符号化の概要と関連研究について述べ，本研究においてこれらの理論をどのように扱うかについて説明する．

1.1.　Schema 理論

Schema とは，発達心理学者である J. Piaget が人の認知発達過程を説明するために提唱した概念であり，人が環境との相互作用を通して獲得する環境や行動に関する認知構造を意味する [1]．言い換えると，Schema は外界の対象や自分の行動の仕方，あるいは行動による環境の変化についての知識を自己組織化して繰り返し利用する枠組みとして考えることができる．

Schema は人が生得的に持っている静的なものではなく，生まれて間もない未分化の状態から，乳幼児が環境との相互作用を繰り返す中で適応的に変化していく動的なものである．Piaget はこのような適応的変化を，環境を適切に認識できるように Schema を変化あるいは分化させる均衡化過程として捉え，生物における不変の機能としている．均衡化は「同化」と「調節」というダイナミクスを通して行われる．同化とは，ある対象あるいは現象が与えられたときに得られる感覚情報をもとに既存の Schema の中から対応するものを同定し，それに基づいて行動を起こすことである．すなわち，自分が持っている知識を使って受け取った感覚情報を理解し，既存の Schema に紐づく感覚運動経験として取り込むことが同化なのである．しかし通常，環境から受け取る感覚情報は過去の経験と同じものではない．乳幼児がこのような新しい経験を同化しようとすると既存の Schema では捉えられなくなる．このとき，新しい経験を理解できるように持っている Schema に修正を加える．これを調節と呼ぶ．Piaget は同化と調節を分けて説明しているが，一方の作用はもう一方の作用を前提としており，実際にはこれらは不可分なものである．

以上のような Schema 理論は乳幼児の発達を説明する上で非常に有力だが，あくまで発達心理学における概念モデルであり，具体的に乳幼児の認知発達がどのような原理で起こるのか，そして Schema の均衡化がどのように進むのかについては表現することができず，直接的な分析が難しい．そこで認知発達メカニズムを計算モデルとして表現し，分析可能な発達主体をつくることを考える．

Schema 理論に基づいた，あるいはそれに類似した計算モデルはこれまでにも提案されている．例えば，川人らは複数の予測・制御モジュールを持ち，制御と学習を行うモジュールを切り替えながら行動する MOSAIC というモデルを提案した [2]．これは，各モジュールの環境に関する予測誤差をもとに算出する責任信号という指標を用いて各時刻において行動や学習を分担する比率を変化させる．利用する認知的な構造を適応的に切り替えるという点で，このモデルは Schema 理論に類似するモデルであるということができる．しかしモジュールの数は事前に設定しており，発達の中で新しいモジュールを獲得することができない．また，谷口らは現在の状態から目標状態を出力するモジュールと，現在の状態から目標状態に移行するための行動を出力するモジュールをそれぞれ知覚シェマ・行為シェマとする双シェマモデルを提案している [3]．このモデルでは各モジュールに対応する記憶領域に感覚運動経験を記憶し，記憶した経験と新しく得た経験それぞれに対する予測誤差に基づいたシェマ活性度という指標を用いてモジュールの切り替えや追加を行う．Schema の分化過程までモデル化しているという点でこのモデルはより Schema 理論に合致していると言えるが，過去の経験を有限の記憶領域にそのまま保持する点や Schema の分化が不連続に行われる点で人の連続的な発達を表現しきることは難しいと考える．そこで本研究では，連続な発達的変化の中で Schema が多様化していく過

程をモデル化するべきであるという考えに基づき，内部モデルの持つ自らの行動や環境の変動によって時間的に変化する内部状態を抽象化した分散表現を Schema として捉えることで，学習や行動の作りこみを抑えた計算モデルの構築を試みた．

1.2.　予測符号化

予測符号化とは人の脳は環境から得られる感覚信号を予測し，誤差を最小化することを基本的な機能としているという考え方である [4]．脳は，環境との相互作用を通して自らが受け取る感覚信号を予測できるように環境や身体に関する内部モデルを学習し，同時に予測誤差を減らすように環境に対して働きかけるように行動決定を行うとされている．この理論は本来，既に発達した脳のはたらきを説明するものだが，予測符号化に基づく乳幼児の発達に関する研究も行われている．S. Kouider et al. は生後 12 か月の乳幼児に対して，聴覚刺激とそれに関連する視覚刺激，あるいは関連しない視覚刺激を提示したときの脳活動を計測した [5]．その結果，遅い知覚情報処理を行う部分は予測外の刺激に対して反応し早い知覚情報処理部分では予測された刺激に対して反応していることが確認され，乳幼児の脳においても感覚予測が行われていることを示唆している．

また，予測符号化に基づいてロボットの行動学習を行う研究も行われている．尾形らは感覚運動情報から近未来の感覚運動情報を予測して実際に受け取る信号との誤差を小さくするように内部モデルを学習する深層予測学習を提案している [6]．この方法で学習したモデルによってロボットは環境から受け取る情報を予測しながら行動を行う．これにより，学習時のデータと行動時のデータに差異がある場合や未学習の環境下においても柔軟に作業を行うことが可能になる．

本研究では, 感覚運動情報の予測符号化が乳幼児の持つ Schema の均衡化を促しているという仮説に基づき， Schema に類似する構造をもちうる内部モデルをニューラルネットワークとして構築して深層予測学習によって学習を行うことで Schema の発達的変化の分析を行う．

2.　予測符号化に基づいた Schema の発達計算モデル

本研究では人の発達プロセスをニューラルネットワークによってモデル化し，近未来に主体が受け取る感覚信号を予測するようにモデルを学習することによって行動を獲得させた．　本章では提案する Schema の発達計算モデルの構造と学習手法について説明する．

2.1.　モデル構造

モデル構造の概要を図 1 に示す．ここでは感覚情報としてカメラからの視覚情報を，運動情報としてロボットの関節角度情報を扱う．モデルは時刻tにおけるカメラ画像s_t^vとロボットの関節角度s_t^pを入力し，それをもとに時刻$t+1$のカメラ画像の予測$\hat{s}_{t+1}^v$及びロボットの関節角度の予測$\hat{s}_{t+1}^p$を出力する．予測符号化に基づく深層予測学習では，モデルの感覚予測値$(\hat{s}_{t+1}^v, \hat{s}_{t+1}^p)$と実際に時刻$t+1$に得られる感覚情報の真値$(s_{t+1}^v, s_{t+1}^p)$との誤差が最小になるように学習を行う．ロボットが行動する際にはカメラ画像と関節角度を取得し，各時刻における関節角度の予測値$\hat{s}_{t+1}^p$に関節を動かし続ける．

モデルは Encoder-Decoder 部とリカレント部を持ち，それぞれ感覚情報－特徴量の変換と特徴量の時間的変化の予測を担う．特に，リカレント部は現在時刻の情報だけでなく前時刻の情報を内部状態として保持し，予測に利用する構造を持つ．Schema が自らの行動や環境の変化という時間変化するものを捉える枠組みであることを考慮すると，このモデルにおいてはリカレント部に Schema が獲得されると考えることが妥当である．そこで，リカレント部の内部状態を時間方向に圧縮した表現を Schema として考える．具体的には，RNN(Recurrent Neural Network)に複数のパターンを埋め込む効果を持つ Parametric Bias(PB)をリカレント部への入力に導入し，学習によって獲得される PB の表現を Schema として扱う．

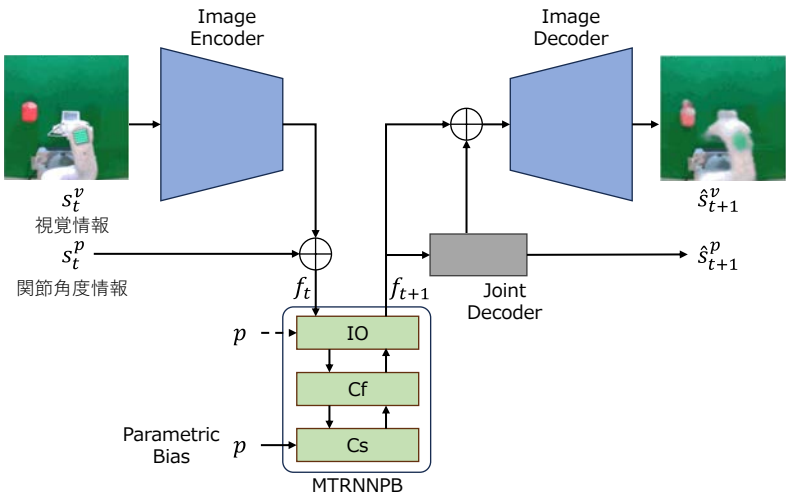

図 1: 予測符号化に基づく Schema 発達モデルの概要

2.1.1. Encoder-Decoder 部

Encoder-Decoder 部では，リカレント部で扱う低次元の特徴量と高次元の感覚信号を変換する．Encoder では感覚信号(s_t^v, s_t^p)から特徴量f_tを抽出する．尚，画像s_t^vからの抽出を行う ImageEncoder には CNN(Convolutional Neural Network)を使用し，関節角度s_t^pは十分低次元であるため圧縮を行わずそのまま特徴量として使用した．すなわち，Encoder 部は次式で表される．

$$f_t = \mathrm{ImageEncoder}(s_t^v) \oplus s_t^p. \tag{1}$$

ここで，$\oplus$はベクトルの結合を表す．Decoder ではリカレント部で予測した次時刻の特徴量f_{t+1}から感覚信号の予測値$(\hat{s}_{t+1}^v, \hat{s}_{t+1}^p)$の再構成を行う．画像$\hat{s}_{t+1}^v$の再構成を行う ImageDecoder には Transposed CNN を用い，関節角度$\hat{s}_{t+1}^p$の再構成を行う JointDecoder には全結合層を使用した．また，ImageDecoder には特徴量f_{t+1}と再構成した関節角度$\hat{s}_{t+1}^p$の両方を入力している．以上から，Decoder 部は次式で表される．

$$\hat{s}_{t+1}^p = \mathrm{JointDecoder}(f_{t+1}), \tag{2}$$

$$\hat{s}_{t+1}^v = \mathrm{ImageDecoder}\left(f_{t+1}, \hat{s}_{t+1}^p\right). \tag{3}$$

2.1.2. リカレント部

リカレント部では，次時刻の特徴量を予測する．ここでは，MTRNNPB を使用した．MTRNNPB は, Multiple Timescale Recurrent Neural Network(MTRNN) [7]と Recurrent Neural Network with Parametric Bias(RNNPB) [8]を組み合わせた構造を持つ．

MTRNN とは，応答速度の異なるニューロン群から構成される階層的な RNN である．MTRNN は入出力層である IO 層とコンテキストの予測を行う Cf 層，Cs 層で構成され，それぞれ再帰的な入力を持つ．時刻tにおけるi番目のニューロンの内部状態$u_{i,t}$と出力$y_{i,t}$は，次式で表される．

$$u_{i,t} = \left(1 - \frac{1}{\tau_i}\right) u_{i,t-1} + \frac{1}{\tau_i}\left[\sum_j w_{ij}\, x_{j,t}\right], \tag{4}$$

$$y_{i,t} = \tanh\left(u_{i,t}\right). \tag{5}$$

ここで，τ_iはi番目のニューロンの時定数，w_{ij}は入力のj番目のニューロンからi番目のニューロンへの結合重み，$x_{j,t}$は時刻tにおけるj番目のニューロンからの入力を表す．式(4)から時定数が大きくなるほど前の時刻の内部状態が保存されるため，内部状態の時間的変化がより圧縮される．これにより時定数が小さいと短期的，大きいと長期的な予測が可能となる．MTRNN では，IO 層，Cf 層，Cs 層の順に時定数を大きく設定することで層の深さに伴って時間的抽象度の高いアトラクタを獲得し，層によ

って時間スケールの異なる予測を行うことができる．

RNNPB とは，RNN に複数のアトラクタを埋め込むために入力に学習可能な定数(Parametric Bias; PB)を追加した RNN である．一般的な RNN では，複数種類の時系列パターンを学習させることが難しく，別々の行動を一つの RNN に埋め込むことができないという問題がある．そこで，RNN の入力に PB 層を導入し，その他の時間変化する入力とは異なり毎ステップ一定の値を入力する．これにより，パターンの違いによって内部状態に強制的に差を生じさせ，パターンの違いに関する情報を陽に与えることで時系列特性の異なるデータを学習させることができる．本研究では，MTRNN が獲得するアトラクタを圧縮したある種の分散表現として PB を用いる．そこで，学習初期に全パターンに対する PB 値を**0**に初期化し，学習によって PB を更新していくことによりパターンに応じて自己組織化させた．PB の更新はネットワーク重みの更新と同時に行われる．PB の更新式は次の式で与えられる．

$$\Delta p_n = k_{bp} \sum_{t=1}^{L_n} \delta_{p_n,t}, \tag{6}$$

$$p_n \leftarrow p_n - \Delta p_n. \tag{7}$$

ここで，p_nはn番目のデータに対する PB 値，L_nはn番目のデータのステップ数，$\delta_{p_n,t}$は時刻tにおける損失関数のp_nに関する勾配，k_{bp}は PB の学習率を表す．

本研究では MTRNN の IO 層あるいは Cs 層に PB を導入したものをリカレント部として用い，Encoder の出力f_tを IO 層に入力し，IO 層の出力をリカレント部の出力f_{t+1}とする．そして，学習によって PB を導入した層に埋め込まれる行動別のアトラクタに対応した PB が獲得されるか，そして学習過程において PB がどのように変化するかを検証する．

2.2.　データ収集と学習

モデルの学習はロボットを動かして収集した感覚運動データを用いて行う．データの収集はロボットの動きを事前にプログラムで作成し実際に動かす，あるいはコントローラ等でロボットを操縦する等の手法でロボットにタスクを実行させ，その時のセンサ情報の時系列データを記録する．本研究ではカメラ画像s_t^vと関節角度s_t^pを収集するので，n番目のデータD_nは

$$D_n = \{(s_0^v, s_0^p), (s_1^v, s_1^p), \dots, (s_{L_n-1}^v, s_{L_n-1}^p)\} \tag{8}$$

と表され，データセット全体は

$$D = \{(D_1, p_1), (D_2, p_2), \dots, (D_N, p_N)\} \tag{9}$$

と表される．ここで，Nはデータの総数を表し，L_nはn番目のデータのステップ数，p_nはn番目のデータに対する PB を表す．尚，p_nは学習前に**0**に初期化され，学習によって変化する点に注意されたい．

データセットDが与えられたとき，学習によって全てのデータ，ステップに関して(s_{t-1}^v, s_{t-1}^p)をモデルに入力したときの予測値$(\hat{s}_t^v, \hat{s}_t^p)$と真値$(s_t^v, s_t^p)$との誤差が最小となるようなネットワークパラメータ$W$と PB 値$\{p_1, p_2, \dots, p_N\}$を求めたい．よって，データ$D_n$に対する損失関数は次式で表される．

$$\mathcal{L}_n = \sum_{t=1}^{L_n} ||\hat{s}_t - s_t||_2^2. \tag{10}$$

ここで，$\hat{s}_t = (\hat{s}_t^v, \hat{s}_t^p)^\top, s_t = (s_t^v, s_t^p)^\top$である．全てのデータにおいてステップ数が一致するとは限らず，データによってステップ数にばらつきが生じる．そこで，学習時には最もステップ数が大きいデータに合わせてほかのデータを時間方向にパディングし，損失を計算する際にはパディング部分をマスクしてデータのある部分のみの損失を計算する．

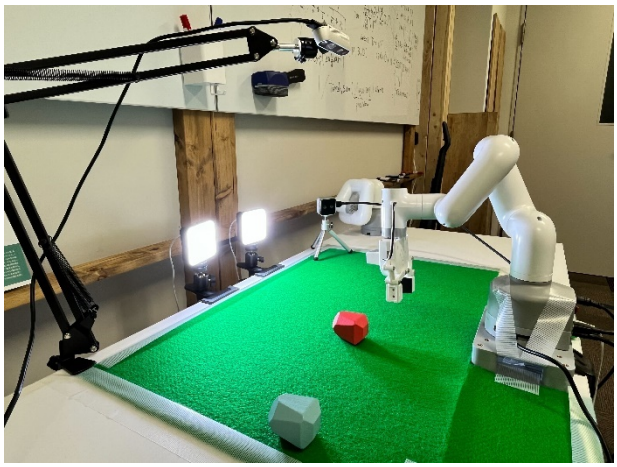

図 2: 実験環境

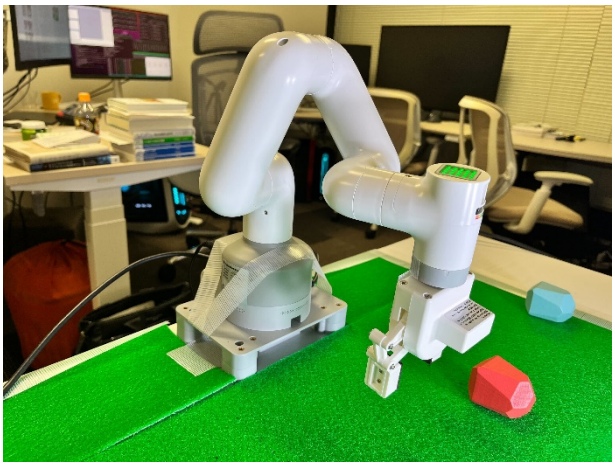

図 3: MyCobot 280 Pi

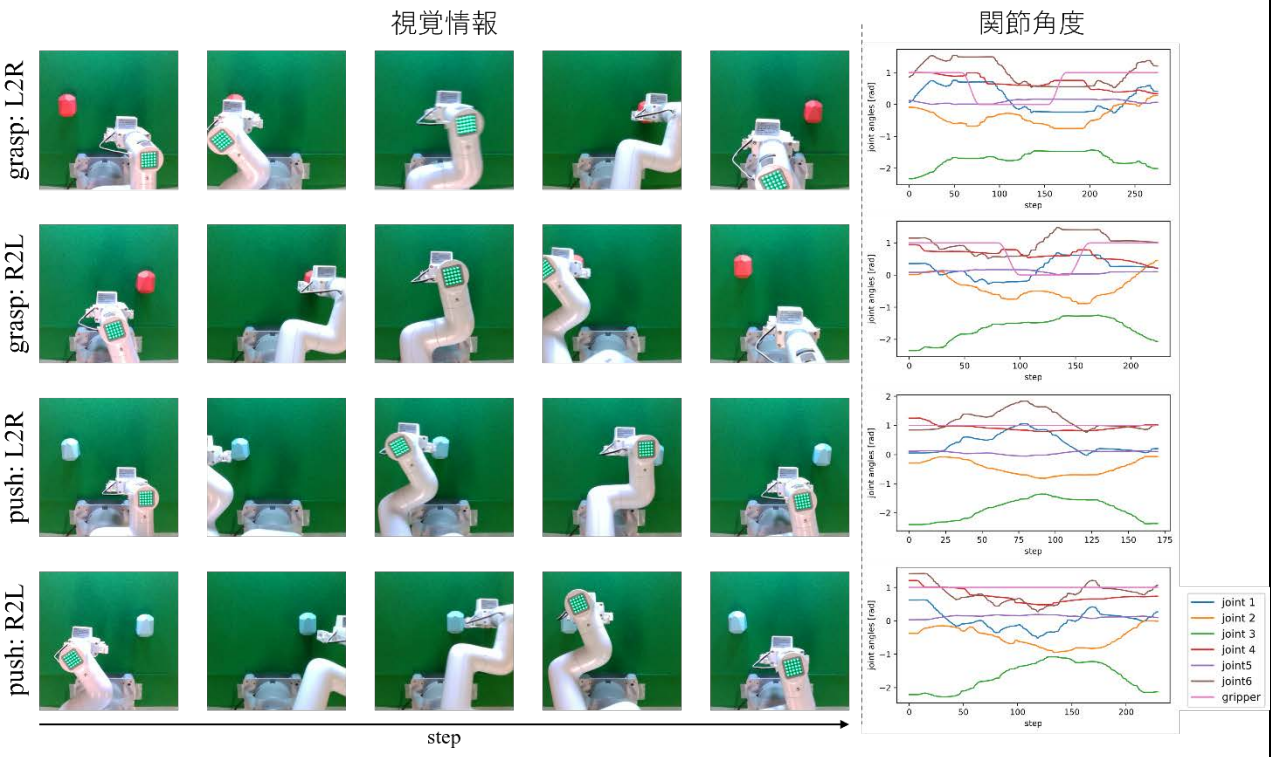

図 4: 作成した感覚運動データセットのサンプル例

3.　実験

人による物体操作タスク実行時に収集した感覚運動情報データを用いてモデルを学習し，行動獲得及び PB の変化について分析を行った．また，学習したモデルを用いて実ロボットを制御し，どのような行動をとるかについて検証した．

3.1.　実験 1; 物体操作行動の学習

実ロボットを操作して物体操作行動のデータを収集し，それを用いてモデルを学習した．そして，学習過程における PB の変化や感覚情報予測について分析した．事前に収集したデータを用いた学習は乳幼児が物体操作行動を教示されて繰り返し行動を試みることを意味し，この実験は乳幼児の行動学習による Schema の発達的変化を分析することに相当する．

3.1.1.　感覚運動情報データセット

図 2 に本研究で使用した環境を示す．尚，行動データ収集時に加えて後述するリアルタイム行動生成実験でも同様の環境を使用した．本研究では，実ロボットとして市販の 6 自由度マニピュレータで

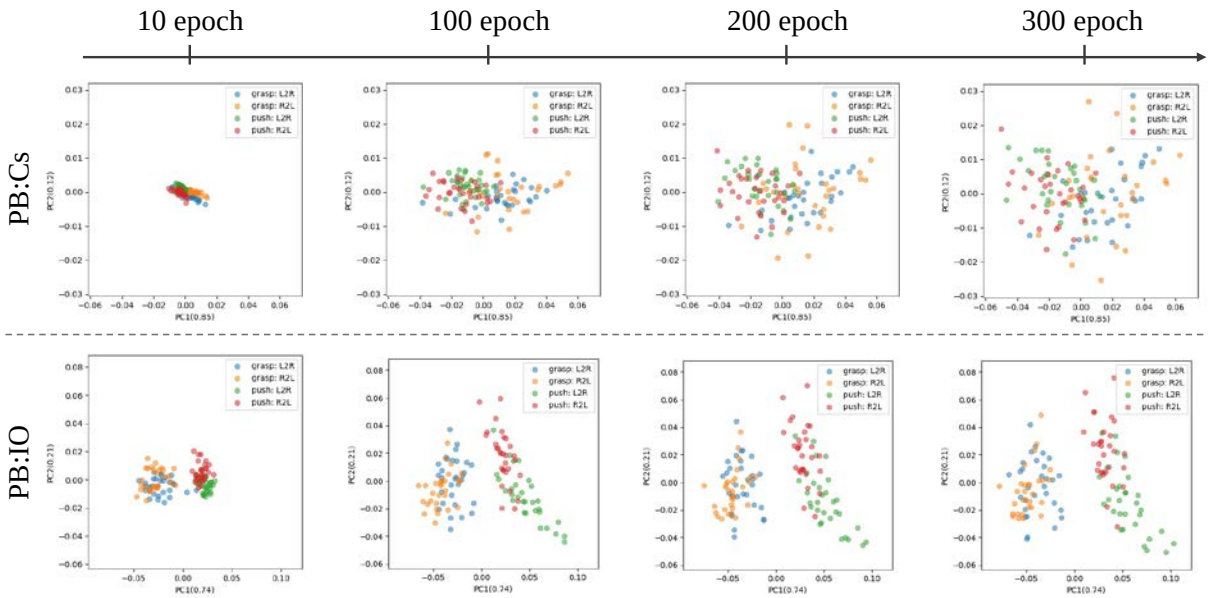

図 5: PB:Cs モデル及び PB:IO モデルが獲得した PB 空間の変化

ある MyCobot 280 Pi (Elephant Robotics)の先端にグリッパ(MyCobot-Adaptive Gripper)を装着したもの（図 3）を使用し，視覚情報を取得するカメラとして Intel RealSense D435 を使用した．ここではゲーム用コントローラを用いてロボットを操縦し，そのときのカメラ画像及びマニピュレータの関節角度を 10Hz のサンプリングレートで記録することによって感覚運動データを収集した．尚，本実験では画像のサイズは$256 \times 256 \times 3$，関節角度情報は 6 自由度とグリッパの開閉情報を合わせた 7 次元とした．

データセットに含まれる感覚運動情報の例を図 4 に示す．収集する行動タスクは，赤色の積み木を把持して左から右あるいは右から左に移動させる行動と水色の積み木をグリッパを開いた状態で押して左から右あるいは右から左に移動させる行動の 4 パターンとした．以降，これらのパターンをそれぞれ”grasp: L2R”，”grasp: R2L”，”push: L2R”，”push: R2L”と呼ぶ．動作の種類（grasp/push）と物体移動方向はそれぞれ積み木の色と初期位置に対応しており，理想的には視覚情報から得られる積み木の色と位置から行動パターンを一意に選択することができる．ここでは，各パターンにつき 30 サンプル，計 120 サンプルの時系列データを収集した．尚，ステップ数L_nの最大値は 308 ステップとなった．

3.1.2.　モデルと学習パラメータの設定

本実験では特徴量f_tの次元数を 32 とし，MTRNNPB の内部状態$u_{i,t}$の次元数を IO 層では 64，Cf 層と Cs 層では 32 に設定した．MTRNNPB の各層の時定数は，試験的に IO 層：2，Cf 層：5，Cs 層：10 とした．また，IO 層あるいは Cs 層に入力する PB の次元数は 4 とした．

また，ネットワークパラメータの学習率を1×10^{-4}，PB の学習率k_{bp}を 0.2 に設定し，300 エポックの学習を行った．尚，最適化アルゴリズムとしてネットワークパラメータの更新には Adam を，PB の更新には確率的勾配降下法を使用した．また，バッチサイズは 1 に設定した．これは，PB を更新する際にミニバッチ内の他のデータの影響を受けないようにするためである．

3.1.3.　結果

Cs 層に PB を導入したモデル（PB:Cs モデル）と IO 層に PB を導入したモデル（PB:IO モデル）の学習過程における PB 空間を図 5 に示す．ここでは，本来 4 次元である PB を主成分分析(PCA)により 2 次元に圧縮している．また，訓練に使用していない評価用のデータを用いて予測した感覚と内部状態を分析した．尚，視覚情報には評価データを入力し，関節角度情報には 0 ステップ目のみ評価

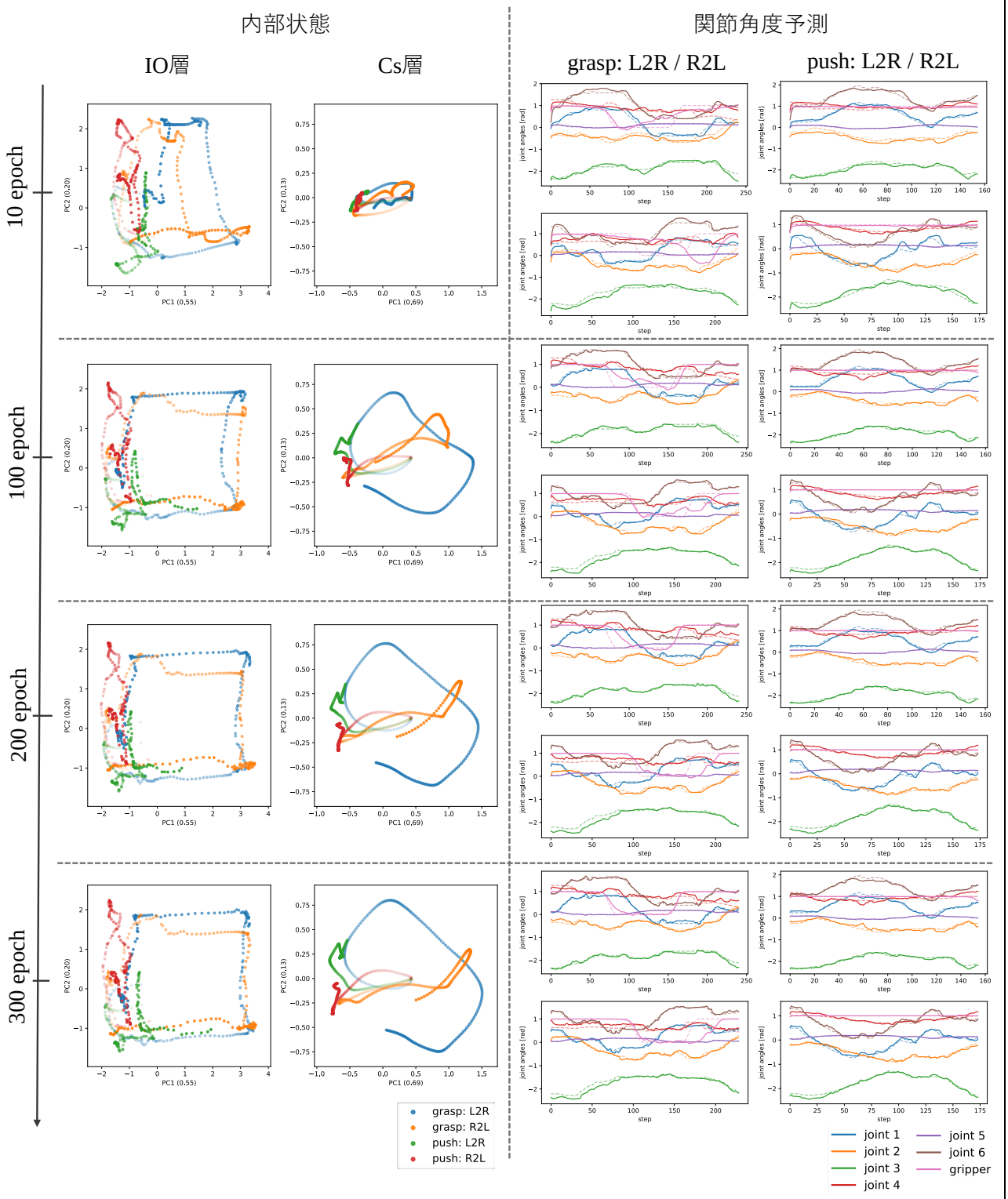

図 6: 評価データに対する，各学習過程における PB:Cs モデルの内部状態及び関節角度予測．内部状態はステップ数が大きい程色を濃くしている．また，関節角度予測では波線が真値，実線がモデルの出力を表す．

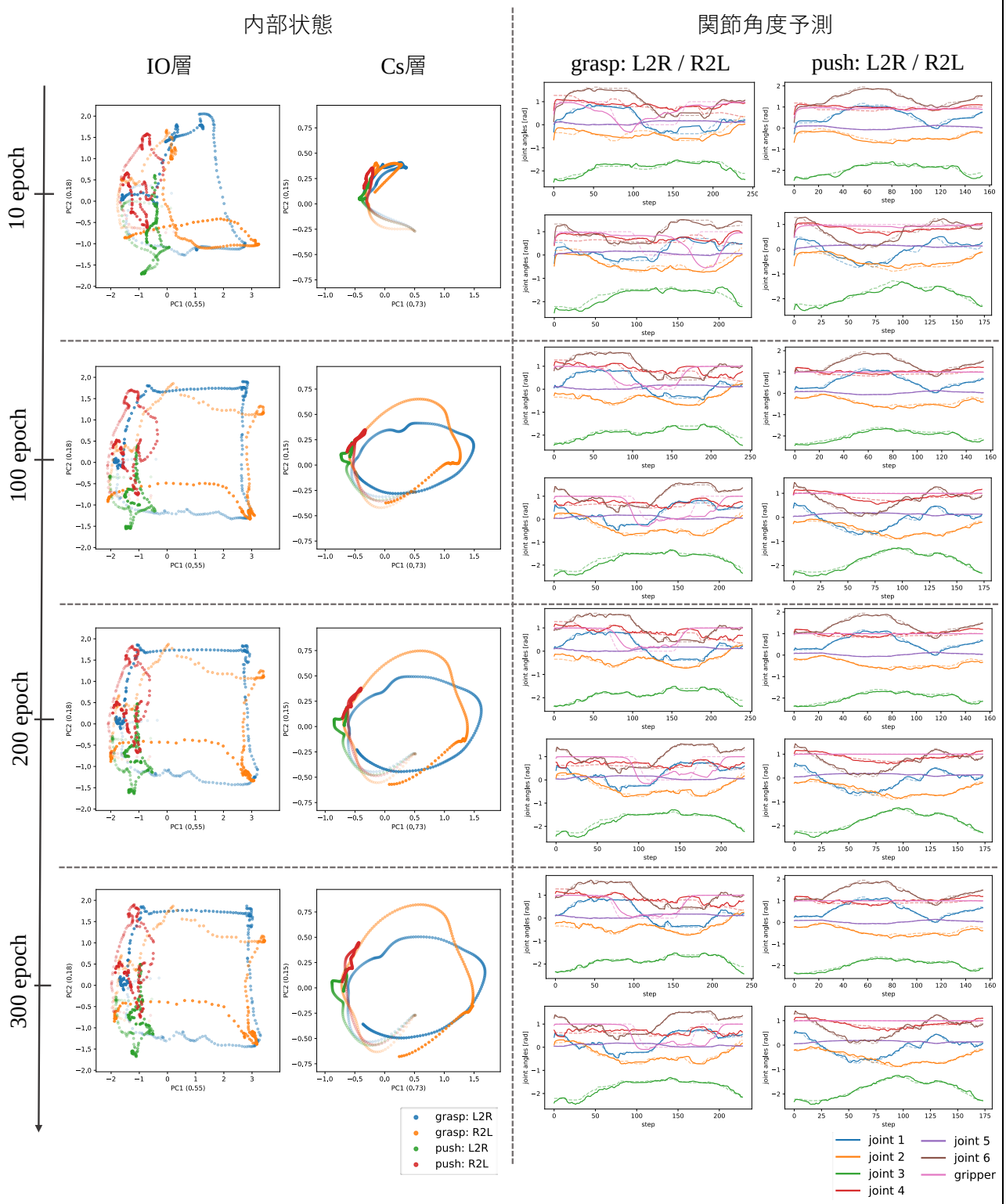

図 7: 評価データに対する，各学習過程における PB:IO モデルの内部状態及び関節角度予測．内部状態はステップ数が大きい程色を濃くしている．また，関節角度予測では波線が真値，実線がモデルの出力を表す．

表 1: 学習済みモデルを用いたリアルタイム行動生成時の各パターンにおける実験結果

	grasp: L2R				grasp: R2L				push: L2R		push: R2L	
	接触	グリッパの開閉	把持	移動	接触	グリッパの開閉	把持	移動	接触	移動	接触	移動
PB:Cs	◯	◯	×	×	◯	×	×	×	◯	×	◯	◯
PB:IO	◯	◯	◯	×	◯	◯	◯	◯	×	×	◯	×

データを，その後はモデルが前ステップに出力したものを入力している．各学習過程のモデルにおける評価データに対する内部状態と関節角度の予測結果を図 6 及び図 7 に示す．尚，内部状態は PCA によって 2 次元に圧縮し，ステップ数が大きい程色を濃くしている．

3.2. 実験 2; 学習済みモデルを用いたリアルタイム行動生成

実験 1 で 300epoch まで学習したモデルを用いて実ロボットを制御し，実環境においてどのような行動をとるか検証した．具体的には，カメラから画像を取得し，取得した画像と直前にモデルが出力した関節角度をモデルに入力してモデルが出力した関節角度にロボットを動かすというループを繰り返すことで実環境の中でロボットに行動させた．尚，モデルに入力する PB は実験 1 で 300epoch まで学習したモデルが獲得した PB の各パターンについての平均とした．また，カメラから画像を取得して関節角度を送信する頻度はデータ収集時と同じく 10Hz とした．

3.2.1. 結果

各行動パターンにおけるリアルタイム行動生成時の結果を表 1 に示す．grasp 行動についてはグリッパが物体に接触したか，グリッパの開閉行動がみられたか，物体を把持したか，物体移動タスクが遂行されたかの 4 段階，push 行動についてはグリッパが物体に接触したか，物体移動タスクが遂行されたかの 2 段階で評価した．

4. 考察

4.1. 実験 1; 物体操作行動の学習

図 5 より PB:Cs モデルと PB:IO モデルに共通して，学習初期には(0.0, 0.0)周辺に密集していた PB が，学習が進むにつれて広がっていることがわかる．PB:Cs モデルにおける PB の変化を図 5 について確認すると，明確な分化は見られないものの徐々に PC1=0 付近を境界として左側には push，右側には grasp の行動に対応する PB が分布するように変化している．一方で，いずれの物体操作方法においても L2R と R2L に対応する PB は混ざり合って分布した．また，PC1 軸の寄与率が 0.85 と大きな値であることから，この PB 空間の持つ情報のほとんどが PC1 軸で表現されている．このような PB 空間が獲得された要因としては，MTRNN の Cs 層において感覚運動情報の短期的な変化，ここでは物体を移動させる方向に関する情報が圧縮され，グリッパの開閉状態といった操作方法に関連する情報のみが残ったためであると考えられる．実際に，図 6 で Cs 層の内部状態を見ると，移動方向に関しては明確な軸がみられないのに対して操作方法に関しては左右方向に表現されていることが確認できる．

対して PB:IO モデルにおける PB の変化を見ると，10epoch の時点で grasp と push それぞれのクラスタが形成され，100epoch 以降では push の中で移動方向に関する PC2 軸への分化が顕著にみられた．さらに図 7 で IO 層の内部状態を確認すると PC1 軸がグリッパの開閉情報を，PC2 軸がグリッパの位置を表現しており，PB:IO モデルでは行動パターンの違いによる内部状態の差が PB にも反映された．

以上の結果から，MTRNN の IO 層に導入した PB では時間方向に関して具体的な認識が，Cs 層に導入した PB では抽象的な認識が行われることがわかった．つまり，IO 層では物体操作の方法が同

じでも移動する方向が異なっている場合は異なる行動として認識される一方，Cs 層では移動方向に関わらず操作方法によって行動を認識する．Schema の観点から捉えると，学習によって大小それぞれの粒度で Schema が形成され，MTRNN 全体で Schema が持つ意味の階層構造を獲得したと考えることができる．

また，学習によってモデルが獲得した行動に注目すると，図 6,7 からどちらのモデルにおいてもグリッパの開閉以外に関しては 10epoch 時点で概ね真値に近い値を出力しており，早い段階からグリッパを物体に近づける行動を獲得していることがわかる．一方で，グリッパの開閉に関しては習得が遅く，grasp:R2L では 300epoch 学習したモデルにおいてもグリッパを開くタイミングが遅れたままとなっている．これは，グリッパを物体位置に移動させるとロボット自らによって物体が隠れ，視覚情報からグリッパと物体の位置関係が分かりにくくなっているためと考えられる．

4.2.　実験 2; 学習済みモデルを用いたリアルタイム行動生成

表 1 から，PB:Cs モデルでは全ての行動パターンにおいて，PB:IO モデルでは push: L2R 以外の行動パターンにおいて物体とグリッパを接触させることに成功している．つまり，本実験で学習したモデルはグリッパを物体位置まで移動させる行動を獲得しているといえる．さらに，PB:Cs モデルでは push: R2L で物体の移動に成功し，PB:IO モデルでは grasp: L2R と grasp: R2L の両方において物体の把持に成功した．以上の結果から，モデルは視覚情報から積み木の色と位置を認識し，学習したパターンの中から行動を選択することができると考えられる．一方で，物体の移動に成功したのは各モデル 1 パターンのみであった．物体位置にグリッパを移動させることはできても物体の移動ができない要因としては，本実験の設定ではグリッパを物体の位置まで移動させるとカメラと物体の間にロボットが位置することになり，物体との位置関係や物体の色が認識できなくなるためであると考えられる．

5.　まとめと今後の展望

本研究では，分析可能な行動主体としての計算モデルを実現するために人の認知発達を説明する概念モデルである Schema と脳の基本的機能とされる予測符号化に注目し，MTRNNPB を用いた Schema 発達計算モデルの構築，深層予測学習による行動獲得実験及び実ロボットを用いたリアルタイム行動生成実験を行った．実験の結果，次時刻の感覚運動情報の予測誤差を最小化するようにモデルを学習することにより，MTRNN の各層に導入した PB の空間に時間的抽象度の異なる Schema が獲得され，MTRNNPB 全体に Schema の持つ意味が構造化された知識が自己組織化されることが確認された．

一方で，本研究では事前に収集した感覚運動データを用いたオフライン学習によってモデルを学習したが，実際の乳幼児は実世界とインタラクションをしながら常に新しい情報を収集して学習していると考えられるため，本研究で用いた手法は完全に人の発達メカニズムをモデル化できているとはいえない．また，学習時や行動生成時にモデルに入力する PB はモデル自身が選択しているわけではなく，学習によって得られた PB をもとに外から与えているに過ぎない．Piaget が提唱する本来の Schema の同化プロセスでは主体が得た感覚運動情報から自ら同化する Schema を選択するはずであり，本研究の学習プロセスは Schema 理論を完全に踏襲できているとは言い難いと考える．よって，今後は感覚運動情報をもとにモデルに入力する PB を生成し，環境とのインタラクションの中で学習を行うことを目指す．

参考文献

[1] J. Piaget, “La naissance de l'intelligence chez l'enfant, ” 1936 (谷村, 浜田 訳: 知能の誕生, ミネルヴァ書房).

[2] M. Haruno, D. M. Wolpert , M. Kawato, “Multiple Paired Forward-Inverse Models for Human Motor Learning and Control,” 1999.

[3] 谷口忠大 , 椹木哲夫, “双シェマモデル　自律エージェントの為の自己組織化機械学習手法の提案,” 人工知能学会論文誌, 19 巻 6 号, 2004.

[4] K. Friston, K. James , H. Lee, “A free energy principle for the brain,” J. Physiology Paris, Vol. 100, pp.70-87, 2006.

[5] S. Kouider et al., “Neural dynamics of prediction and surprise in infants,” Nature Communication, vol.6, 2015.

[6] H. Ito, K. Yamamoto, H. Mori , T. Ogata, “Efficient multitask learning with an embodied predictive model for door opening and entry with whole-body control,” ScienceRobotics,Vol.7, Issue65, 2022.

[7] Y. Yamashita , J. Tani, “Emergence of Functional Hierarchy in a Multiple Timescale Neural Network Model: A Humanoid Robot Experiment, ” PLoS Computational Biology, vol. 4, no. 11, 2008.

[8] J. Tani , M. 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, 2003.

謝辞

本研究の遂行にあたり, ご多忙の中ご指導をいただいたアドバイザー教員の堀井隆斗講師に深くお礼申し上げます.

様式６

申請先学部　　基礎工学部　採択番号　No.7