

# ミメシス原理にもとづく運動認識・生成から知能への接近

## - ミラーニューロンの発達モデル -

稲邑 哲也 (JST CREST/東大)

中村 仁彦 (東大)

An approach from motion generation/recognition to intelligence

based on the mimesis principle

- A developmental model of mirror neurons -

\*Tetsunari INAMURA (JST CREST/Univ. of Tokyo),

Yoshihiko NAKAMURA (Univ. of Tokyo)

**Abstract**— The discovery of Mirror Neurons in human brain shows that the motion recognition and generation are executed by bidirectional computation model, and the model leads to symbol grounding system through embodied humanoids. We have proposed mirror neuron models based on Hidden Markov Models. In this paper, we propose a new method for usual models in order to be developed through behavior observation and generation.

**Key Words:** Imitation Learning, Mirror Neurons, Hidden Markov Models, Symbol Emergence

### 1. はじめに

ヒューマノイドの全身行動の設計手法として、人間の動作を真似て新しい行動を獲得するアプローチが多く提案されている<sup>1)2)</sup>。しかしながら、ただ単純に行動を見まねするだけでは、運動パターンの表現を他者の座標系から自己の座標系に変換しているだけに過ぎない。人間が行なっている見まねは、単なる座標変換ではなく、相手の行動を認識し、それを自分の意志として実行するというプロセスである。このためには、まず他人の行動を観測を通じてシンボルとして認識し、その後シンボルから自分の行動を再現する、という2つのプロセスが必要となる。我々はこの行動認知と行動生成の相互的な仕組みの中にヒューマノイドの知性の発現のキーが存在していると考えている。

行動認知と行動生成の関係については、認知心理学の分野でも近年注目が集まってきており、ミメシス<sup>3)</sup>という枠組が提案されている。ヒトの脳は300万年前に現在のヒトと同等の容積に達しているのに対し、音節言語は5万~30万年前程度の比較的新しい機能である。この事実は音声言語によるコミュニケーション以前に身振りによるシンボルコミュニケーションが成立していた事を示唆する。つまりミメシスは、行動認知と行動生成の循環によってシンボル操作能力を得たとする仮説原理である。

一方、生物学の分野からも興味深い事実が発見されている。他人の特定の行動を観測する時に発火し、自分がその行動を行なう時にも発火するミラーニューロンがサル脳の言語野において発見されている<sup>4)</sup>。この事から、ミラーニューロンは他人の行動をシンボルとして抽象化した表現形態であると解釈でき、また同時に、ミラーニューロンが認識プロセスと逆方向の生成プロセスとして働いた結果、自分の行動が生じると解釈する事が可能である<sup>5)</sup>。

これらの事実は行動の認識と生成、すなわち見まね

を行うミメシスの機能モデルが脳のハードウェアに埋め込まれており、シンボル操作などの高次知能の基礎となっていることを示唆する。著者らは、これらのパラダイムにヒントを得て、シンボルを獲得、操作して、全身で行動するヒューマノイドの知能の構成論の研究を進めている。

我々は行動の認識と生成を同時に学習し、シンボル表現の創発を同時に行なう数学モデルとして、隠れマルコフモデル (Hidden Markov Models, 以下 HMM) を用いた手法を提案してきた<sup>6)7)8)</sup>。しかしながら従来のモデルは、行動を表現する基本単位である行動要素のや隠れマルコフモデルの明確な設計指針が存在せず、観察した行動に適応できなかつたり、自分自身の身体性と整合性の合わない行動が生成されたりする、などの問題があった。本論文では、他人の行動を認識し、自己の行動を生成する、という繰り返しを行ないながら後天的にミメシスモデルを発達させる枠組を提案する。また、ヒューマノイドにおける実験を通じて、ミラーニューロンの発達モデルの実用性を示す。

### 2. 運動認識 / 生成の双方向計算モデル

#### 2.1 ミラーニューロンモデルの意義

本章では今までに提案されてきた見まね学習システムと、我々が目指すミラーニューロンモデルの相違点について説明しながら、システムの概要を示す。

鮫島らによって提案された MOSAIC による見まねの枠組<sup>9)</sup>では、運動パターンを予測し制御するための順動力学と逆動力学モジュールが複数存在し、適切なモジュールの担当割合を制御することで観測した運動パターンを真似ることができる。しかしながら、この枠組は参照されるべき手本となる運動パターンの存在が前提となっており、時系列情報が含まれないシンボル表現から運動パターンを作り出すことが難しい側面がある。

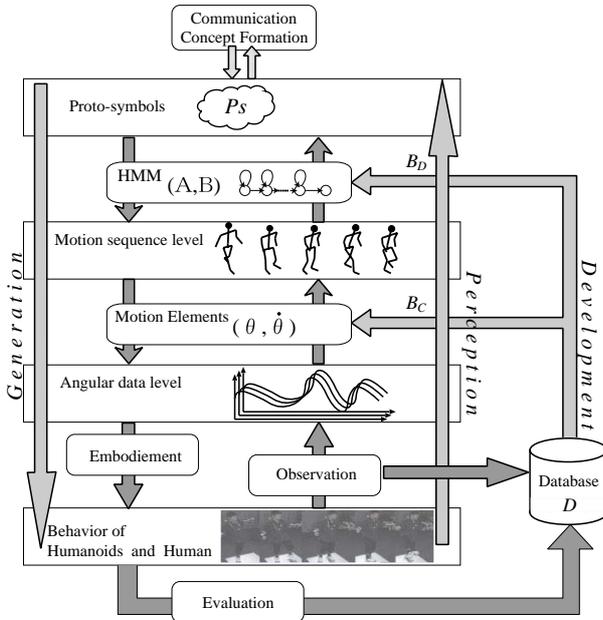


Fig.1 A proposed mirror neuron model

また、國吉らの提案した teaching by showing<sup>10)</sup> と呼ばれる見まねの枠組では、ある程度複雑な行動が抽象化モジュールで表現されており、行動を再現することは容易である。しかしながら、刻一刻と変化して行く運動パターンのダイナミクスの抽象化という観点の考察はなされておらず、行動の移り変わりはほぼシンボル表現のまま記述されて行く。

そこで、我々は、全身行動に含まれるダイナミクスをシンボルとして抽象化し、逆にシンボルを用いて行動の再現と行動の認識を行なう手法を提案する。もし、行動のダイナミクスを抽象化し、シンボルとして表現することができれば、先に述べたミラーニューロンを工学的に構成することに等しい。

## 2.2 ミラーニューロンモデルの概要

本研究で提案する行動認識・獲得システムの概要を Fig.1 に示す。システムはまず、人間やヒューマノイドの行動を観察し、その行動微小時間ごとに区切って行動要素 (Motion Elements) に変換する。行動要素とは微小時間内でのヒューマノイドロボットの各関節角度や角速度、トルクなど低レベルの物理量で表現される。観察された他者の行動、および生成する自己の行動は行動要素の時系列データとして表現され、この時系列データを抽象化することによってシンボル表現を獲得する。本研究ではこのモデルによって獲得されるシンボルの元となる概念を「原始シンボル」(Proto-symbol)と呼ぶ。一方、抽象化とは逆方向のプロセスで行動を生成することが可能である<sup>11)</sup>。

## 2.3 HMM による双方向計算モデル

観察された他者の行動は行動要素の時系列データとして表現され、原始シンボルはこの時系列データを抽象化することによって得られる。我々はこの抽象化のために、時刻データを認識するモデリング手法として知られている HMM を採用した。

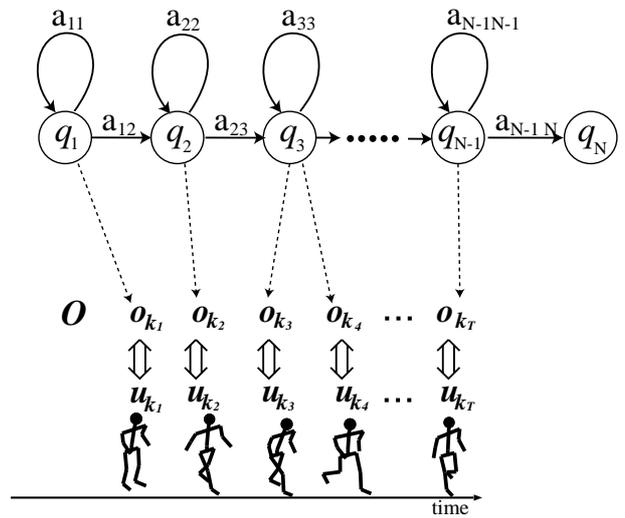


Fig.2 Motion Sequence and Hidden Markov Models

HMM は状態の有限集合  $Q = \{q_1, \dots, q_N\}$ 、出力ベクトルの有限集合  $S = \{o_1, \dots, o_M\}$ 、状態遷移確率分布  $A = \{a_{ij}\}$ 、記号出力確率分布  $B = \{b_{ij}\}$ 、初期状態確率分布  $\pi = \{\pi_i\}$ 。以上の5項組  $\lambda = \{Q, S, A, B, \pi\}$  で表される。Fig.2 のように、HMM は確率的に状態遷移を繰り返しながら、ベクトルを出力する。このベクトルは離散的なラベル表現をとる事もあり、その場合は離散型 HMM、それ以外の通常モデルを連続分布型 HMM と呼ぶ。本研究では前者の離散型 HMM を採用し、出力ラベルを行動要素と一対一で対応付ける。 $a_{ij}$  は状態  $i$  から  $j$  への遷移確率を示し、 $b_{ij}$  は状態  $i$  からベクトル  $o_j$  を出力する確率である。このように HMM は状態遷移を行いながらデータを出力するダイナミクスを表現する確率モデルであるため、運動パターンを認識するだけでなく、生成するという観点から見ても親和性が高い。以上が HMM を採用する理由である。

HMM のパラメータ  $\lambda$  は、あるカテゴリに属する複数の行動要素列  $O = \{o_{k_1}, o_{k_2}, \dots, o_{k_T}\} (k_i \in \{1, 2, \dots, N\})$  が最も良く出力されるようなパラメータが設定されている。すなわち、 $\lambda$  が行動のカテゴリを代表するパラメータであると見なせる。本研究では Fig.2 に示すように左型のノードから右側のノードへと状態遷移が進んでいく Left-to-Right 型の HMM が使用されているため、初期ノードは必ず  $q_1$  であり、 $\pi$  は考慮する必要がない。また、HMM の構造は各行動で共通であるため、 $Q, S$  も行動カテゴリを代表するパラメータに含む必要はない。よって式 (1) のように  $A, B$  の組が原始シンボル  $\mathcal{P}_S$  と定義される。この原始シンボルは一つの行動カテゴリにつき一つづつ割り当てられる。

$$\mathcal{P}_S \stackrel{\text{def}}{=} \{A, B\} \quad (1)$$

ヒューマノイドの全身行動は有限個数の行動要素で表現されるため、行動要素は関節角度やトルクで構成される位相空間上の点  $\mu$  に相当する。有限個数の点で位相空間をカバーするために、各要素にはその点を中心とする担当範囲を示す共分散行列  $\Sigma$  が付加される。す

なわち，行動要素  $u_i$  は次式のように定義される．

$$u_i \stackrel{\text{def}}{=} \{\mu_i, \Sigma_i\} \quad (2)$$

ここで，

$$o_i \stackrel{\text{def}}{=} u_i \quad (3)$$

の定義を導入する事で，HMM による全身行動の表現を行なう．本研究では行動要素の物理量として関節角度を扱うこととする．

### 3. HMM による行動の記憶・認識・再生

#### 3.1 他者行動の記憶と原始シンボル生成

観測された行動パターンを行動要素列に変換するために，まず微小時間で全身行動をサンプリングし，その値  $\theta$  に対して

$$j = \arg \max_i \frac{1}{\sqrt{(2\pi)^n |\Sigma_i|}} \exp\left\{-\frac{1}{2}(\theta - \mu_i)' \Sigma_i^{-1} (\theta - \mu_i)\right\} \quad (4)$$

の計算を行ない，担当する  $j$  番目の行動要素  $u_j$  を対応付け，行動要素列  $O = \{o_{k_1}, o_{k_2}, \dots, o_{k_T}\}$  を求める．ただし  $n$  は要素表現の次元数である．

次に行動要素列からこれを尤もよく発生させるような HMM のパラメータ  $A, B$  を獲得し，原始シンボルとして登録する．パラメータ推定には EM アルゴリズムの一種である Baum-Welch アルゴリズム<sup>12)</sup> を用いる（付録参照）．

#### 3.2 HMM による他者行動の認識

観測された行動パターンの認識には，行動要素列  $O$  が HMM (すなわち原始シンボル) から出力される尤度  $P(O|A, B)$  を Viterbi アルゴリズム<sup>12)</sup> を用いて計算する．入力された行動パターンに対応する原始シンボルは高い尤度を示し，入力パターンと関係のない原始シンボルは低い尤度を示すはずである．そこで認識判定指数

$$R(O) = \log \frac{\max\{P(O|P_{S_i})\}}{\text{second}\{P(O|P_{S_i})\}} \quad (5)$$

を導入する．ただし， $\text{second}\{x\}$  は，集合  $x$  のうち，2 番目に大きい値を意味するものとする．この値が大きければ，他の原始シンボルと混乱することなく，他者の行動を認識したとみなし，

$$j = \arg \max_i \{P(O|P_{S_i})\} \quad (6)$$

の  $j$  を用いて  $P_{S_j}$  を認識結果とする．認識判定指数  $R$  の値が低ければ，認識は失敗，すなわち未知の行動と見なして，原始シンボルの生成フェーズに移る．

#### 3.3 原始シンボルを用いた自己の全身行動の生成

HMM はパターン認識の分野で広く用いられている手法であるが，本研究では逆に HMM から行動パターンを生成する手法を導入する．

尤度  $P(O|P_S)$  の計算を行なう事で，ある行動要素列に対応する HMM を選択し認識する事は可能であるが，逆に原始シンボル  $P_S$  から行動を生成するには， $P(O|P_S)$  が最大となる  $O$  を求める事になり，一般的

な解法が存在しない．音声合成の研究分野において，連続分布型 HMM を用いてパターン生成する例<sup>13)</sup> も見られるが，生成される特徴ベクトルの種類に制限があるため，本論文では一般的な特徴ベクトルに対しても生成が可能な手法を採用する．

HMM からのパターンの生成は，パターン  $O$  が取り得る空間全体に対して探索を行い，もっとも相応しいと思われるパターンを見つけ出すことと同等であるが，行動要素列の長さを  $l$  とした場合，パターンの取り得る組合せは  $M^l$  と膨大なものになり，最尤法を適用する事もできない．そこで遺伝的アルゴリズム (GA) を用いた探索法を導入する．HMM から出力される行動要素列の長さはあらかじめ一定の長さ  $l$  を取ると仮定し，一つの遺伝子を一つの行動要素と対応させ， $l$  個の遺伝子で  $O$  を表現する．個体に対する適合度として，尤度  $P(O|A, B)$  を用いる．

最終的に，行動要素の時系列パターンを実際の関節角度のパターンに変換し，ヒューマノイド上で実現可能な行動を出力する．

### 4. 行動の生成・評価のループによるミラーニューロンモデルの発達

HMM の性質を決定するパラメータと行動要素は EM アルゴリズムの一種である Baum-Welch アルゴリズム<sup>12)</sup> で求められる．このアルゴリズムは時系列の行動データの集合であるデータベース

$$D = \{O^1, O^2, \dots, O^r\} \quad (7)$$

を入力とし，以下の操作

$$\{A, B\} := B_D(D) \quad (8)$$

$$\{\theta, \Sigma\} := B_C(D) \quad (9)$$

で獲得される．ただし， $B_D, B_C$  は Baum-Welch アルゴリズムによるパラメータ決定の操作を表す<sup>7)</sup>．

しかしここで，二つの問題が生じる．一つ目の問題は，初期のデータベースの状態は教示者の行動のみが含まれており，この状態で行動要素や原始シンボルを決定すると，学習者の身体特性とは異なる性質が抽出されてしまう点である．二つ目の問題は，残りの HMM パラメータ，特に状態の有限集合  $Q$  を決めるノードの個数  $N$  を決定する指針が与えられていないことがある．そこで，行動の認識と生成を繰り返しデータベースを操作することによって，適切な原始シンボルと行動要素を他者行動の認識と自己行動の生成にとも有用となるように発達させる．具体的な手順は以下のようになる．

1. 原始シンボル  $P_S$  と行動要素群から行動  $O$  を生成
2. 生成した行動  $O$  が適切なものかどうかを判定
3. 適切ならば，データベースに生成した行動を追加． $D^{t+\Delta t} := \{D^t, O\}$
4. 原始シンボルと行動要素を上記の式 (8)(9) で再学習し，1 に戻る

2) の評価の基準として，自らが獲得した原始シンボルの質を評価するための内部評価と，大局的な行動の意味などを教示者の立場から評価するための外部評価，

Table 1 Recognition result of others' motion using HMMs.

Input Behavior	Proto-symbols						R
	swinging	walking	dancing	kicking	backward walking	crawling	
swing	-430	-3915	-4077	-3940	-4114	-4007	3485
walking	-3048	-225	-3071	-1646	-3099	-3019	1420
dance	-1656	-1603	-144	-1613	-1683	-1577	1433
kicking	-2543	-1574	-2562	-199	-2585	-2519	1374
backward walking	-2395	-2318	-2413	-2332	-202	-2372	2117
crawling	-4083	-3950	-3815	-3976	-4151	-488	3327
unknown behavior	-1915	-1853	-1928	-1865	-1946	-1896	11

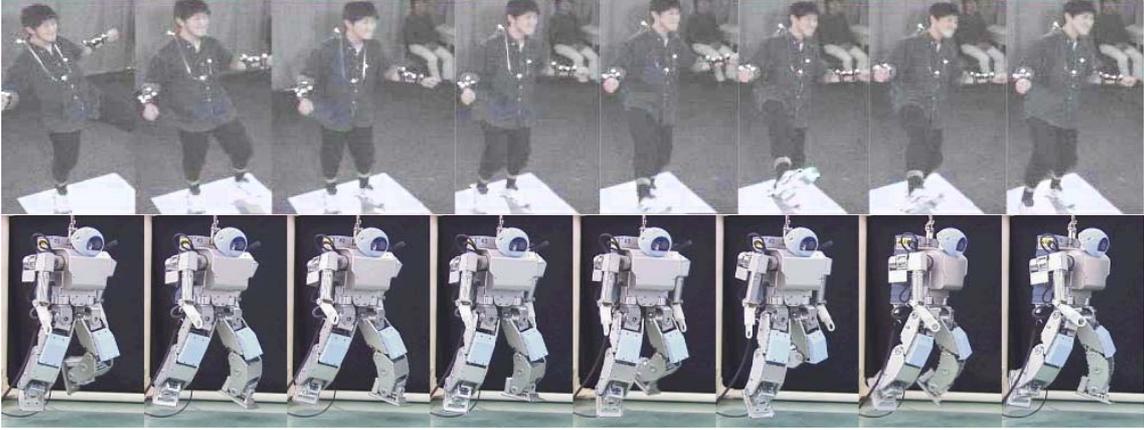


Fig.3 Original humans' performance (upper) and generated humanoid's behavior (lower)

の二つの評価基準を用意した．本論文では内部評価として式 (5) で導入した認識判定指数  $R$ ，外部評価として，生成した行動と観察した行動の関節角度の誤差の時間積分

$$E_{\theta} = \frac{1}{T} \int_0^T |\theta_{in}(t) - \theta_{out}(t)| dt \quad (10)$$

を用いた．ただし， $\theta_{in}(t)$  は外部評価として与えられた理想的な運動の関節角度系列， $\theta_{out}(t)$  は実際に再生された運動の関節角度系列， $T$  は生成された運動の時間を示す．この二つの評価基準をまとめて

$$V = \alpha E_{\theta} + \beta R^{-1}(O) \quad (11)$$

を用いて評価を行ない，ある閾値を越えた場合には自己の身体特性に沿った適切な運動であると認識し，データベースに追加する．それ以外の場合には運動データは棄却される．ここで  $\alpha, \beta$  は評価のバランスを統括する定数である．

## 5. ヒューマノイドを用いた行動認識・生成・発達実験

実験に用いるヒューマノイドは(株)富士通製，小型ヒューマノイド HOAP-1 である．片腕 4 自由度，片足 6 自由度，合計で全身に 20 自由度を持つ．リアルタイムに全身の関節角を計測できるモーションキャプチャシステム<sup>14)</sup>を用いて，人間の行動を観察し，見まねを介した原始シンボル生成の有効性を検証した．この実験では 20 自由度の身体を持つヒューマノイドが対象となるので，行動要素も 20 次元のベクトルとなる．

### 5.1 行動認識実験

歩行(walk)，後ろ歩き(backwalk)，キック(kick)，ハイハイ歩き(haihai)，テニスのスイング(swing)，コサックダンス(dance)の 6 種類の行動をあらかじめ学習させ，学習データには用いていない未知データを加えた 7 種類の行動データを用いて行動の認識実験を行なった．その結果を Table.1 に示す．

表の数値は尤度  $P(O|A, B)$  の対数である．縦に並んでいる行動の名前が認識対象となる行動  $O$  で，横に並んでいる行動の名前が既に学習してある原始シンボル  $P_S (= \{A, B\})$  に相当する．この値が大きければ原始シンボルに良くマッチした行動であると判断できる．ここで，表の対角成分に注目すると良く分かるが，観測された行動と同じ行動に対する原始シンボルの尤度が高くなっている．また，学習されていない未知の行動(unknown)に対する尤度はどれも同じような値を示しており，経験的に識別判定指数  $R$  の値を  $-1000$  と設定しておけば，間違わずに認識できる事が分かった．

### 5.2 行動生成実験

対象となる行動パターンは人間の歩行(その場での足踏み)動作である．第 3.3 節で説明した手法に基づいて生成したロボットの動作パターンとその元となった人間の参照行動を Fig.3 に示す．

図から分かる通り，おおまかではあるが，人間の行動と同等の動作を生成できることが確かめられた．しかしながら，現段階ではヒューマノイドは空中で吊られている状態である．バランスを取るなどの動作スキルに関する学習は小脳で行なわれている説が有力であり，我々が目指しているのはその小脳に入力すべき参照データを生成するミラーニューロンモデルの確立で

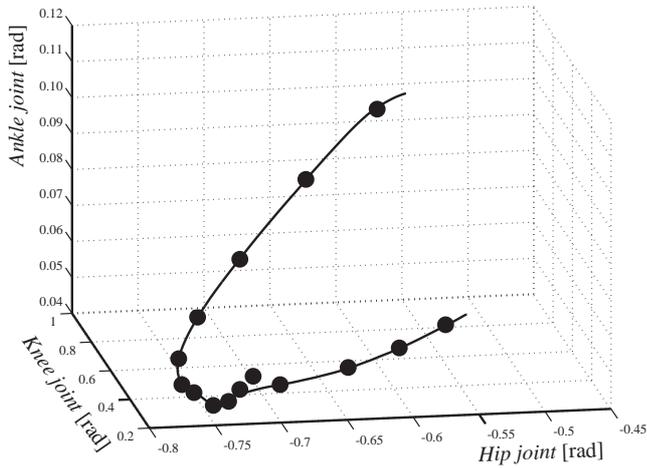


Fig.4 Acquired motion elements without loop structure

あるという立場から、動力学条件の考慮は別の枠組で行なうべきであると考えている。

### 5.3 行動要素の発達実験

ヒューマノイドとは異なる身体特性を持つ人間の行動の見まねから、適切な行動要素が獲得される事を確かめるため、ヒューマノイドの膝関節の稼働範囲は人間のそれよりも小さく、約40度であるという設定を設け、スクワット行動を対象に行動の発達実験を行った。

80回の行動生成を繰り返して作成されたデータベースから獲得された行動要素を Fig.4,5 に示す。ここでは全身の自由度のうち、膝関節および股関節と足首関節のピッチ成分の3自由度が示されている。図中の曲線は観察した行動の軌跡で、各点が行動要素を示している。Fig.4が稼働範囲の制限が無い場合、Fig.5が稼働範囲の制限がある場合の結果である。双方とも、15個の行動要素が獲得されるように学習条件を統一した。

観察した行動の軌跡に近い場所に行動要素が集まってきているが、手本となる行動に対応するAの領域だけでなく、Bで示した領域にも行動要素が分布している。これは、稼働範囲の狭い身体特性に沿うように生成した行動によって得られたものである。すなわち、他者行動を認識するために使用する行動要素(Aの領域)と、自己行動を生成するために使用する行動要素(Bの領域)の双方を獲得することが確認された。

### 5.4 HMMの構造決定実験

次に、HMMのノードの数を決定させるため、4章で説明した評価基準  $E_\theta$  とノードの個数の関係を調べた。Fig.6はノードの個数別に生成された行動の関節角度の軌跡を示している。ノードの数を増やすと共に、生成された行動(実線)が理想の行動(破線)に近づいて行く事が分かる。また、Fig.7は、ノードの個数と  $E_\theta$  の関係を示している。以上の結果から、今回のケースではノード数を25にすれば適切な行動生成が可能なが分かった。

通常HMMの構築の際、一般的なノード数の評価指標は存在しないが、ミラーニューロンの発達モデルは、この問題にも適用する事ができる利点がある。

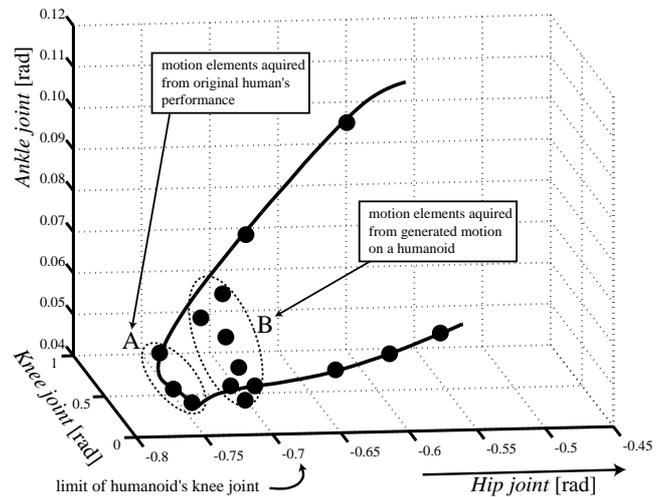


Fig.5 Acquired embodied self-motion elements using loop structure

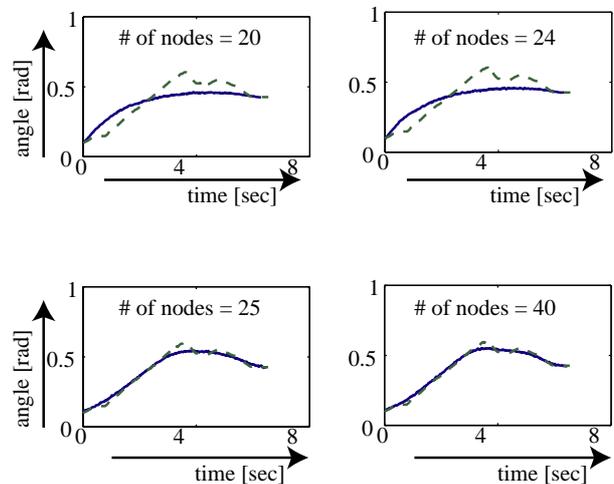


Fig.6 Generated motion (shoulder's yaw joint) for each number of nodes

## 6. 結論

行動の認知、生成からシンボルの創発へとつながるミラーニューロンの工学モデルとしてHMMの持つ数理的性質が有利となる事を示し、実際にヒューマノイドにおける実験を通じてその有効性を確かめた。また、ミラーニューロンモデルの発達の問題を、行動データベースの発達の問題に帰着させるアプローチを提案し、教示者と学習者の身体的特性が異なる場合においても行動要素が適切に獲得される事を確認した。

シンボルが持つ特徴の一つとして、複数のシンボル間に互いの類似関係や距離関係が定義されているという側面がある<sup>15)</sup>。HMMによる原始シンボル表現はパターン情報とシンボル情報の相互変換を可能にするが、ある二つのパターン情報間に存在する行動の意味などの類似性は原始シンボル表現では失われている。現在、行動の類似性を表現可能な原始シンボルへの拡張を行っている段階である。

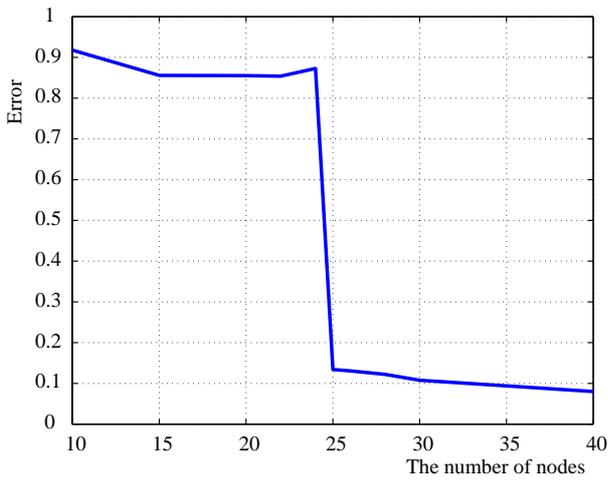


Fig.7 Error value  $E_\theta$  and the number of nodes

なお，本研究は科学技術振興事業団の戦略的基礎研究推進事業「自律的行動単位の力学的結合による脳型情報処理機械の開発」(代表：中村仁彦)の支援を受けた。

#### 参考文献

- 1) M.J. Mataric. Getting humanoids to move and imitate. *IEEE Intelligent Systems*, pp. 18–24, 2000.
- 2) S. Schaal. Is imitation learning the way to humanoid robots? *Trends in Cognitive Sciences*, Vol. 3, No. 6, pp. 233–242, 1999.
- 3) M. Donald. *Origins of the Modern Mind*. Harvard University Press, Cambridge, 1991.
- 4) V. Gallese and A. Goldman. Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, Vol. 2, No. 12, pp. 493–501, 1998.
- 5) 小嶋. ミラーニューロンと言語の起源. *岩波科学*, Vol. 69, No. 4, pp. 404–408, 1999.
- 6) T.Inamura, Y.Nakamura, H.Ezaki, and I.Toshima. Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop. In *the Proc. of IEEE Int'l Conf. on Robotics & Automation*, pp. 4208–4213, 2001.
- 7) T.Inamura, I.Toshima, and Y.Nakamura. Acquisition and embodiment of motion elements in closed mimesis loop. In *the Proc. of IEEE Int'l Conf. on Robotics & Automation*, pp. 1539–1544, 2002.
- 8) 稲邑, 中村. 教示者と学習者の身体差を吸収するミラーニューロンモデル. 第 20 回日本ロボット学会学術講演会予稿集, p. 3H18, 2002.
- 9) 鮫島, 片桐, 銅谷, 川人. 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. *電子情報通信学会論文誌*, Vol. J85-D-II, No. 1, pp. 90–100, 2002.
- 10) Y.Kuniyoshi, M.Inaba, and H.Inoue. Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance. *IEEE Transaction on Robotics and Automation*, Vol. 10, No. 6, pp. 799–822, 1994.
- 11) 稲邑, 戸嶋, 江崎, 中村. ミメシスループと原始シンボルを用いた全身行動の生成. 第 18 回日本ロボット学会学術講演会予稿集, pp. 801–802, 2000.
- 12) S.Young et al. *The HTK Book*. Microsoft Corporation, 2000.
- 13) 徳田, 益子, 小林, 今井. HMM からの音声パラメータ生成アルゴリズム. *信学技報 SP95-122*, pp. 35–42, 1996.

- 14) K.Kurihara, S.Hoshino, K.Yamane, and Y.Nakamura. Optical motion capture system with pan-tilt camera tracking and realtime data processing. In *the Proc. of IEEE Int'l Conf. on Robotics & Automation*, pp. 1241–1248, 2002.
- 15) T. W. Deacon. *The symbolic species*. W.W. Norton & Company. Inc., 1997.

#### 付録

##### 原始シンボルの獲得アルゴリズム

原始シンボルを構成するパラメータ  $A, B$  を求めるアルゴリズム  $B_D$  はまず,

$$\gamma_t(i, j) = \frac{\alpha_t(i)a_{ij}b_j(o_{t+1})\beta_{t+1}(j)}{\sum_{i=1}^N \alpha_T(i)} \quad (12)$$

$$\gamma_t(i) = \sum_{j=1}^N \gamma_t(i, j) \quad (13)$$

を定義し，次式の EM アルゴリズムによって新しいパラメータの推定を行なう．

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \gamma_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad (14)$$

$$\hat{b}_{i(k)} = \frac{\sum_{t:o_t=k} \gamma_t(i)}{\sum_{t=1}^T \gamma_t(i)} \quad (15)$$

この後に，次式でパラメータの更新を行ない，収束するまで式 (14)(15) の推定を続ける．

$$a_{ij} = \hat{a}_{ij} \quad (16)$$

$$b_{i(k)} = \hat{b}_{i(k)} \quad (17)$$

##### 行動要素の獲得アルゴリズム

行動要素を構成するパラメータ， $\mu$  と  $\Sigma$  を求めるアルゴリズム  $B_C$  は以下の EM アルゴリズムで実現される．

$$\hat{\mu}_j = \frac{\sum_{t=1}^T L_j(t) \mathbf{o}_t}{\sum_{t=1}^T L_j(t)} \quad (18)$$

$$\hat{\Sigma}_j = \frac{\sum_{t=1}^T L_j(t) (\mathbf{o}_t - \hat{\mu}_j) (\mathbf{o}_t - \hat{\mu}_j)'}{\sum_{t=1}^T L_j(t)} \quad (19)$$

ただし

$$L_j(t) = \frac{1}{P(\mathbf{O}|A, B)} \alpha_j(t) \beta_j(t) \quad (20)$$

$$\alpha_j(t) = \left\{ \sum_{i=2}^{N-1} \alpha_i(t-1) a_{ij} \right\} b_j(o_t) \quad (21)$$

$$\beta_i(t) = \sum_{j=2}^{N-1} a_{ij} b_j(o_{t+1}) \beta_j(t+1) \quad (22)$$

ここで，初期条件は次式

$$\alpha(1) = 1 \quad (23)$$

$$\beta(1) = \sum_{j=2}^{N-1} a_{1j} b_j(o_1) \beta_j(1). \quad (24)$$

とし，収束するまで式 (18)(19) の推定を繰り返す．