

〈一般研究課題〉 分人型ヒューマン-ロボットインタラクション
システムの開発と次世代サービスロボットへの応用
助成研究者 愛知県立大学 小林 邦和



分人型ヒューマン-ロボットインタラクションシステムの 開発と次世代サービスロボットへの応用

小林 邦和
(愛知県立大学)

Development of a Dividual-based Human-Robot Interaction System and Its Application to the Next Generation Service Robot

Kunikazu Kobayashi
(Aichi Prefectural University)

Abstract :

Currently, in the field of human-robot interaction (HRI), robots have a problem that can only interact the same at all times with humans. In this paper, therefore, we introduce both the concept called a dividual and a general idea of shaping uses for behavior analysis, and build a model of the dividual to grow through interactions with others. In addition, by using a modular neural network and reinforcement learning (Actor-Critic, in particular), we confirmed processes to choose an appropriate dividual out of plural dividuals.

1. はじめに

ロボットは近年、医療や災害の現場で活躍するなど、幅広い分野で研究開発がおこなわれている。そして、知らずしらずのうちにロボットは我々の生活の場に入り込んできている[1-3]。例を挙げれば、スマートフォンに搭載された人工知能(AI)技術による音声エージェントロボット、大手通信会社が販売している人型コミュニケーションロボットなどがある。日本でも、安倍首相は「ロボット革命」を経済成長戦略の目玉としている[4]。このように世間ではロボットに注目が集まっている中、ヒューマン・ロボット・インタラクション(HRI)では、ロボットが人間に対して画一的な対応しかできないという問題を抱えている。しかし、人間は、人に応じて対応を変えることができ、多様なインタラクション(コミュニケーションなど)を実現している。これまでの人間とロ

ロボットのインタラクションは、人間が事前に決められた動作や発話をおこなうことにより、ロボットに一方的に命令を与え、あるタスクを実行するものが多かった[5-9]。しかしこれでは、人間に対して決め細やかな支援をし、人間と日常生活を共にできるロボットの実現には至らない。

インタラクションに関係の深い研究事例として、ロボットに人間のような心理を持たせることで社会的関係の構築を図る研究がある。小嶋は、社会的インタラクションを通してコミュニケーション能力を獲得していく発達モデルに関する研究をおこなっている[10]。また、人間社会における役割をロボットに演じさせることで社会的関係を構築する研究がある。柴田は、アザラシ型ロボットを開発し、ユーザがロボットに触れたり、撫でるといった身体的なインタラクションを通して楽しみや心理的な安寧を得ることを報告している[11]。また、納谷らも同様に、人間-ロボット間の触覚インタラクションに関する研究をおこなっている[12]。しかし、これらの先行研究は、多様な局面での人間とのインタラクションとコミュニケーションを実現する高度な社会性、自律性、適応性の獲得には至っていない。

そこで本論文では、HRIの分野において、文学者の平野啓一郎が提唱した分人の概念と行動分析学におけるシェイピングの概念を導入することにより、他者とのインタラクションを通じて成長する分人モデルの構築を目指す。そこで、我々はモジュール型ニューラルネットワークと強化学習(Actor-Critic)の手法を用いて、分人モデルを構築する。入力情報として、予め分人の形成に必要な情報(カテゴリ情報と行動情報)を定義することにより、複数ある分人から適切な分人が選択される過程を確認する。

以降2節では、平野啓一郎が提唱した分人の概念について紹介する。3節では、行動分析学におけるシェイピングの概念について述べる。4節では、分人モデルについて解説する。5節では、分人モデルに与える入力情報の設計について説明し、計算機シミュレーションとその結果の考察をおこなう。6節では、本論文のまとめと今後の展開について述べる。

2. 分人

分人(dividual)とは、環境や対人関係により、異なる「自分」になるという概念である[13-16]。分人は、相手との反復的なコミュニケーションを通じて、自分の中に形成されていく人格であり、対人関係ごとに定義される。また、相手が一方的に、こうだと決めつけて演じるものではなく、相手との相互作用の中で生じるものである。分人には、図1のように3つのパターンがあるものとする。



図1：分人の概念図

1つ目は、社会的な分人である。これは、知らない人やなじみの薄い人とインタラクションをするための標準的な分人である。2つ目は、グループ向けの分人である。これは、学校のクラスやテニスクラブのような特定のグループのための分人である。3つ目は、特定の相手に向けた分人である。これは、家族や親友のような特定の人のための分人である。本論文では、簡単化のため最後の分人、つまり特定の相手に向けた分人を扱う。

3. シェイピング

シェイピング (shaping) とは、やらせたい行動に近い行動を強化しながら、少しずつ目標としている行動に近づけていくという概念である[17-19]。

行動分析学では、行動が強化されるときに必ず正の強化と弱化、負の強化と弱化が存在する。また、行動を強化する刺激や条件などで学習者にとって「愉快的なものを「好子」、不快なものを「嫌子」という用語を用いて、行動の強化が説明される。そして、行動分析学の根本である人間の行動法則の「好子出現の強化」と「嫌子消失の弱化」が成り立つ。シェイピングは学習者が容易に実行できる行動から複雑な行動へ段階的、誘導的に強化信号を与え、次第に希望の行動系列を形成していく。シェイピングでは初めに目標行動を決める。分人モデルにおける最終目標は、ロボットが分人を理解してインタラクションをおこなうという行動である。次に、最終目標を達成するために一連の行動を分化強化/分化弱化していく。図2に示すように、(a)は好子出現の強化におけるシェイピング

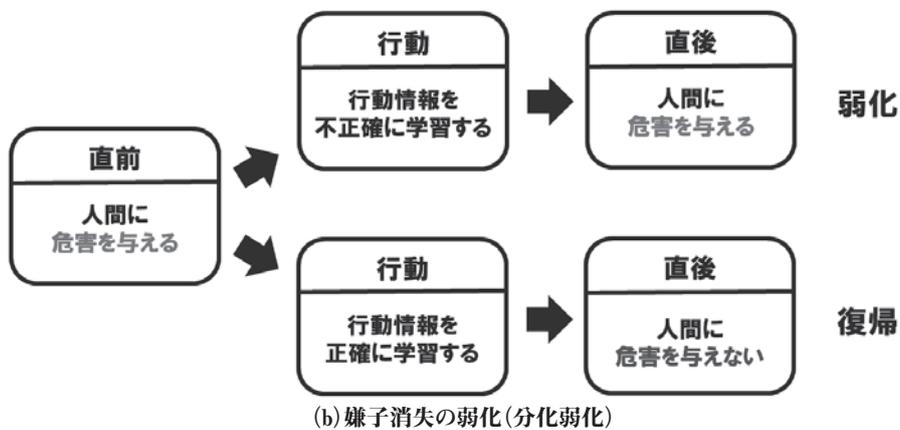
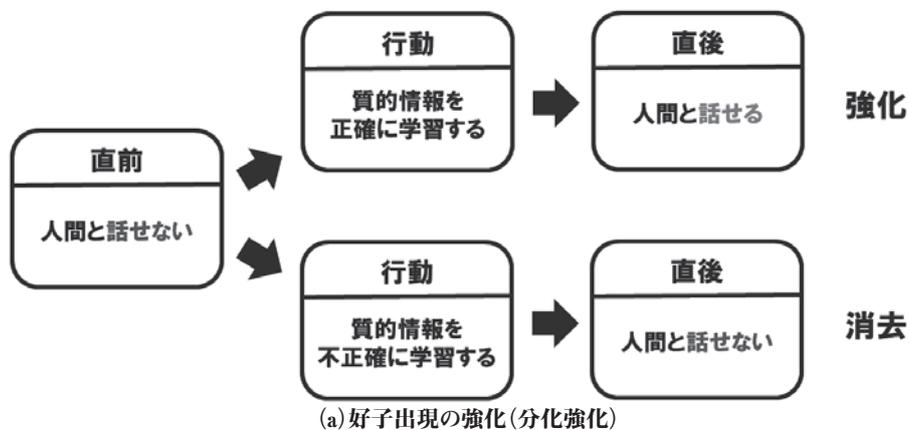


図2：2種類のシェイピング

(分化強化), (b)は嫌子消失の弱化におけるシェイピング(分化弱化)を示している。ここで, 分化強化/分化弱化とは, 特定の行動は強化/弱化し, それ以外の行動は消去する(間接的に弱化する)/復帰させる(間接的に強化する)という手続きである。

本論文では, 生物の学習メカニズムは社会性を構築していくロボットの学習において相性が良いと考え, 行動分析学におけるシェイピングの概念を強化学習に取り入れる。そして, 人間が学習者を見て, 望ましい行動を取った時点でシェイピングをおこなうことを考える。

4. 分人モデル

家族との対話を通して成長し, 学習していくような新たなHRIシステムを想定し, 我々は, 分人モデル[20]の構築をおこなう。

4.1 分人モデルの設計

本論文では, 分人モデルの大枠としてモジュール型ニューラルネットワークを導入する[21]。図3に示すように, 1つのモジュールが1つの分人に対応する構成である。モジュール型ニューラルネットワークを利用する理由としては, 分人のネットワークには中心の人格が存在しない点にある。このことを表現するため, 本論文では, モジュール型ニューラルネットワークを用いることとした。

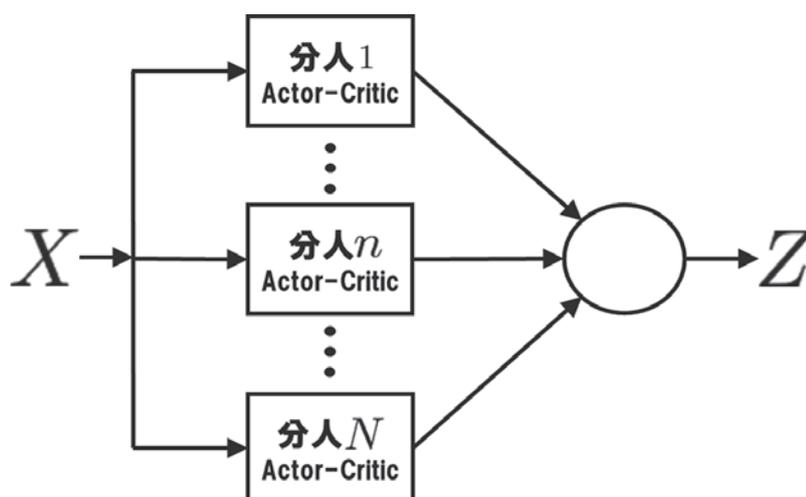


図3: 分人モデルの構成

はじめに, 分人モデルは入力情報を与え学習することによって, モジュールとして分人を形成する。これはつまり, 人間とロボットとのインタラクションの部分に相当する。次に, 各モジュールから出力される値をもとに, 入力情報に対する適切な分人を選択する。これは, ロボットが人間に対してインタラクションを変えられることに相当する。

4.2 分人モデルの学習

各モジュールを学習するために, 図4に示すActor-Critic型強化学習の手法を導入する[22]。ここで, 強化学習とは, 環境との相互作用の結果に対する評価, すなわち報酬の期待値が最大化されるように, 望ましい出力やその時間系列を試行錯誤的に獲得する学習の枠組みのことである。その代

表的な手法の一つがActor-Criticである。Actor-Criticは、行動選択をするActorと呼ばれる部分と、Actorで選択された行動を評価するCriticと呼ばれる部分から構成される。

ここで、強化学習のActor-Critic手法を利用する理由としては、分人が自分で勝手に生み出す人格ではなく、常に環境や対人関係の中で形成される点にある。そして、そのスイッチングは中心の司令塔が意識的におこなっているものではなく、相手次第でオートマチックになされている。これら表現するために、本論文では、Actor-Critic手法を適用したニューラルネットワークを用いることとした。

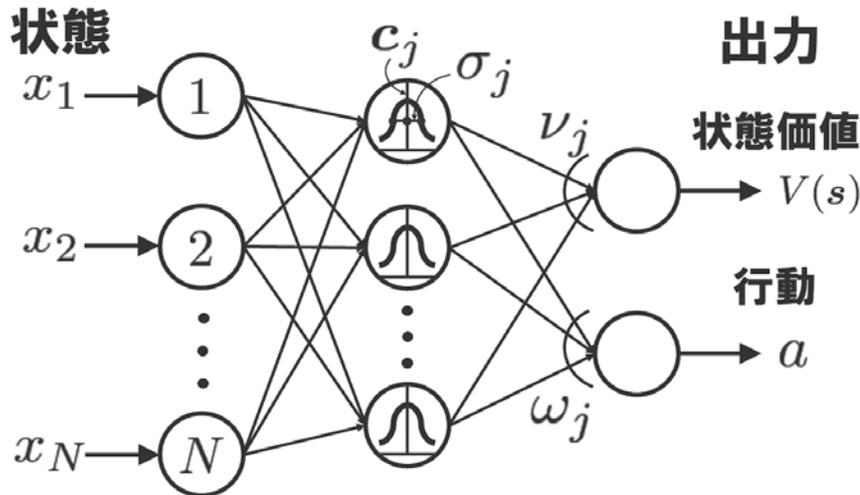


図4: Actor-Critic型ニューラルネットワークの構成

各パラメータは以下の数式で、繰り返し修正される。

- TD誤差

Criticは報酬 r と次の状態 $\mathbf{s}' = (x_1, x_2, \dots, x_N)^T$ を観測し、CriticとActorの学習指標となるTD誤差(Temporal Difference Error)を計算する。ある時刻でのTD誤差 δ は式(1)のように表される。

$$\delta = r + \gamma V(\mathbf{s}') - V(\mathbf{s}) \quad (1)$$

ここに、シェイピングの要素を組み込む。更新式を式(2)のように変形し、新たにシェイピング報酬 r_s を定義する[23,24]。シェイピング報酬とは、従来の予め環境に設定された報酬とは異なり、人間が適宜与えるサブ報酬と定義する。

$$\delta = r + r_s + \gamma V(\mathbf{s}') - V(\mathbf{s}) \quad (2)$$

ここで、 γ は割引率と呼ばれ、 $0 \leq \gamma \leq 1$ の係数である。

- 中間ユニットの学習

中間層の出力関数は、式(3)で表される平均(中心値) c_j 、分散(標準偏差) σ_j のガウス関数 y_j として扱う。

$$y_j = \exp\left(-\frac{\|\mathbf{s} - \mathbf{c}_j\|^2}{2\sigma_j^2}\right) \quad (3)$$

入力値に対して、中心値を近づけていくように式(4)を用いて学習していく。

$$\mathbf{c}_j \leftarrow \mathbf{c}_j + \zeta \delta \nu_j \frac{\mathbf{s} - \mathbf{c}_j}{\sigma_j^2} y_j \quad (4)$$

ここで、 ζ は学習係数と呼ばれ、 $0 \leq \zeta \leq 1$ の係数である。また、ガウス関数の分散 σ_j^2 については、簡単化のため固定している。

- Criticの学習

TD誤差 δ を零に近づけるように、つまり正確な状態価値の予測値を出力するために、式(5)で重み ν_j の学習をおこなう。

$$\nu_j \leftarrow \nu_j + \eta \delta y_j \quad (5)$$

ここで、 η は学習係数と呼ばれ、 $0 \leq \eta \leq 1$ の係数である。

- Actorの学習

状態価値 $V(\mathbf{s}_t)$ が高くなるように、式(6)で重み ω_j の学習をおこなう。

$$\omega_j \leftarrow \omega_j + \rho \delta \quad (6)$$

ここで、 ρ は学習係数と呼ばれ、 $0 \leq \rho \leq 1$ の係数である。

- モジュールの選択確率

式(7)は、Gibbs分布によるソフトマックス手法である。これにより、設計された入力情報が適切なモジュールへ入力されるように学習をおこなう。

$$\exp(p(\mathbf{s}')) p(\mathbf{s}) = \frac{\exp(q(\mathbf{s})/T)}{\sum_{\mathbf{s}' \in X} \exp(q(\mathbf{s}')/T)} \quad (7)$$

ここで、 X は入力情報の集合を表し、パラメータ $q(\mathbf{s})$ は式(8)によって更新される。

$$q(\mathbf{s}) \leftarrow q(\mathbf{s}) + \beta \delta \quad (8)$$

ここで、 β は学習係数と呼ばれ、 $0 \leq \beta \leq 1$ の係数である。また、 T は温度定数である。

5. 計算機シミュレーション

計算機シミュレーションを通して、我々は、複数ある分人の中から適切な分人が選択される過程を確認する。

5.1 入力情報の設計

本論文では、入力情報として、予め分人形成をおこなうために必要な情報を定義しておく。主に情報には2種類あるものとする。1つ目は、個人を識別する「カテゴリ情報」、例えば、名前、性別、年齢、国籍、趣味などである。2つ目は、インタラクションを表現する「行動情報」、例えば、問いかけ、返答などである。識別の都合上、行動情報部分の入力次元は少なくしておく。入力情報の設計にあたり、ニューラルネットワークでの学習を実現するために、以下の操作を順におこなう。

- ラベル付け

入力情報に対する要素のラベル付けを手順1および2によっておこなう。これにより、個人情報データを数値化する。

- 手順1

ラベル付けされた入力情報を2進数に変換する。ただし、「-1」の箇所はビット数分「-1」で埋めることとする。これにより、予め設計したビット数分の情報量をもたせる。

- 手順2

「0」を「-1」、 「1」は「1」、 「-1」を「0」に変換する。これはニューラルネットワークで重みの更新を考慮したとき、「0」であると何も変化しないことを意味するためにおこなう操作である。

このように入力情報は、「-1」、「0」、「1」の値をもつベクトル情報として与えられるものとする。また、情報系列の順序については、人間の場合、情報を得るのはランダムであると考えられる。しかし、ある時刻で見れば、全ての情報は既に揃っていて、それらを特定の順番で入力させていると考えれば不自然ではない。

具体的にはまず、個人の入力情報を表1のように決定する。次に、表2をもとに入力情報に対する要素のラベル付けをおこなう。例えば、ある人の名前がAならば、「1」とラベルを付ける。また、その人の性別が男性ならば、「1」とラベルを付ける。そして、彼の年齢が21歳ならば、「2」とラベルを付ける。ほかの要素も同様にラベル付けをおこなうと、最終的に、「1 1 2 1 2 -1 -1」と表すことができる。ここで、最後の2つの「-1」は行動情報において内容が未確定であることを意味する。次に、表3に示すように入力情報の手順を2回行う。手順1として、ラベル付けしたものを2進数に変換する。これにより、「001 01 10 01 010 -1-1 -1-1」と表すことができる。続いて手順2として、「0」は「-1」、 「-1」は「0」に変換する。

これにより、「-1-11 -11 1-1 -11 -11-1 00 000」と表すことができる。「*」は、行動情報が変化することを表し、インタラクションや分人の成長を表現するためのものである。本来は、「-1」、「0」、「1」の値が入る。

表1: 入力情報

名前	性別	年齢	国籍	趣味
Aさん	男	21歳	アジア	スポーツ
Bさん	男	40歳	アメリカ	読書
Cさん	女	30歳	アジア	料理
Dさん	女	10歳	アジア	ゲーム
Eさん	男	19歳	ヨーロッパ	音楽

表2: 入力集合のラベリング要素

		-1	1	2	3	4	5
カテゴリー	名前 (3ビット)	-	A	B	C	D	E
	性別 (2ビット)	-	男	女	-	-	-
	年齢 (2ビット)	-	1~20	21~30	31~40	-	-
	国籍 (2ビット)	-	アジア	ヨーロッパ	アメリカ	-	-
	趣味 (3ビット)	-	読書	スポーツ	料理	音楽	ゲーム
行動	問いかけ (2ビット)	なし	休みの日はどこに行きますか? etc.			-	-
	返答 (3ビット)	なし	買い物に行くよ! etc.			-	-

表3: 入力情報の変換

個人情報	ラベル付け	手順1	手順2
Aさん	11212 **	001 01 10 01 010 ** ***	-1-11 -11 1-1 -11 -11-1 ** ***
Bさん	21331 **	010 01 10 10 001 ** ***	-11-1 -11 1-1 1-1 -1-11 ** ***
Cさん	32313 **	011 10 11 10 011 ** ***	-111 1-1 11 1-1 -111 ** ***
Dさん	42115 **	010 01 11 10 101 ** ***	-11-1 -11 11 1-1 1-11 ** ***
Eさん	51124 **	010 01 01 10 100 ** ***	-11-1 -11 -11 1-1 1-1-1 ** ***

5.2 シミュレーション設定

本実験では、日常生活での人間とロボットのインタラクションを次のように想定した。ある日、ロボットがある人と接したとき、その彼/彼女に関する情報を得て、ある対応をとる。そして、次の日ロボットが同じ彼/彼女と接するとき、その人のこと考え、行動を変化させるような対応をとる。我々は、以上により、人間とロボットが対等なコミュニケーションをおこなえると考え。それらを考慮し、我々は、1人あたり4パターン、つまりカテゴリー情報は同じで行動情報が違う情報を定義する。

本論文では、分人モデル構築にあたり強化学習を利用する。強化学習では、教師付き学習とは異なり、目標とする出力を学習主体に教えることはしない。学習主体の行動が良かったのか、悪かったのかを評価するだけで(行動に対して報酬や罰を与えるだけで)学習主体自らが試行錯誤のうちに

その目標出力を獲得するという学習をおこなう。また、行動分析学におけるシェイピングの概念を導入したことにより、やらせたい行動に近い行動を強化しながら、少しずつ目標としている行動に近づけていく。そこで、入力情報に応じて、従来の予め学習空間に設定された報酬に加え、新たにシェイピング報酬を設計する。また、表4に示すように人間の行動法則である好子出現の強化および嫌子消失の弱化も考慮し、報酬の設定をおこなう。また、更新式における各パラメータ設定は表5に示す通りである。

表4: シェイピング報酬のパラメータ設定

(a) 好子出現の強化(分化強化)の場合

	従来の報酬	シェイピング報酬
良い行動	15.0	0.0
大変良い行動	15.0	20.0
それら以外の行動	0.0	0.0

(b) 嫌子消失の弱化(分化弱化)

	従来の報酬	シェイピング報酬
良い行動	0.0	0.0
大変良い行動	0.0	0.0
それら以外の行動	-15.0	-20.0

表5: 更新式のパラメータ設定

パラメータ	値
TD誤差の割引率	$\gamma = 0.99$
中間ユニットの学習係数	$\zeta = 0.001$
分散	$\sigma^2 = 1.0$
criticの学習係数	$\eta = 0.01$
actorの学習係数	$\rho = 0.01$
ソフトマックス手法の学習係数	$\beta = 0.01$

5.3 シミュレーション結果

図5と図6はそれぞれ、表4に示した好子出現の強化と嫌子消失の弱化の両者を考慮したシェイピング報酬の設定で、3人分の入力情報を与えたときの各学習ステップにおける累積報酬と累積シェイピング報酬をプロットしたグラフである。両者共に学習回数を重ねるにつれて減少と収束の傾向が見られる。累積報酬の収束値が負の値であるのは、好子出現の強化ではなく嫌子消失の弱化がおこなわれていることを示している。そして、累積シェイピング報酬の収束値が負の値であるのは、分化強化に比べ分化弱化が高い割合でおこなわれたことを示している。つまり、分人モデル内で

は、表4に示した(b)の嫌子消失の弱化におけるシェイピング報酬の設定が適応されており、不適切な学習の機会を分化弱化によって削っていることになる。結果的に、嫌子消失の弱化における負の報酬値を獲得していくことになり、その累積値は負の値となる。後述する適切な分人の選択(図7)を踏まえれば、嫌子消失の弱化によるシェイピングが正確におこなわれていることが読み取れる。本来の強化学習における累積報酬の最大化という目的に相反し、減少の傾向が見られるという結果になっている点については、シェイピング報酬の設定の仕方や与え方および各更新式のパラメータの設定について検討の余地がある。また、収束の傾向が見られることから、各モジュールが正確に分人を学習していることがわかる。つまり、分人の形成がおこなわれていることを意味している。以上より、行動分析学における人間の2つの行動法則とシェイピングの概念を導入したことにより、人間と分人モデルとの関わりなどの外界からのインタラクションの一部が設計できていることが確認できる。

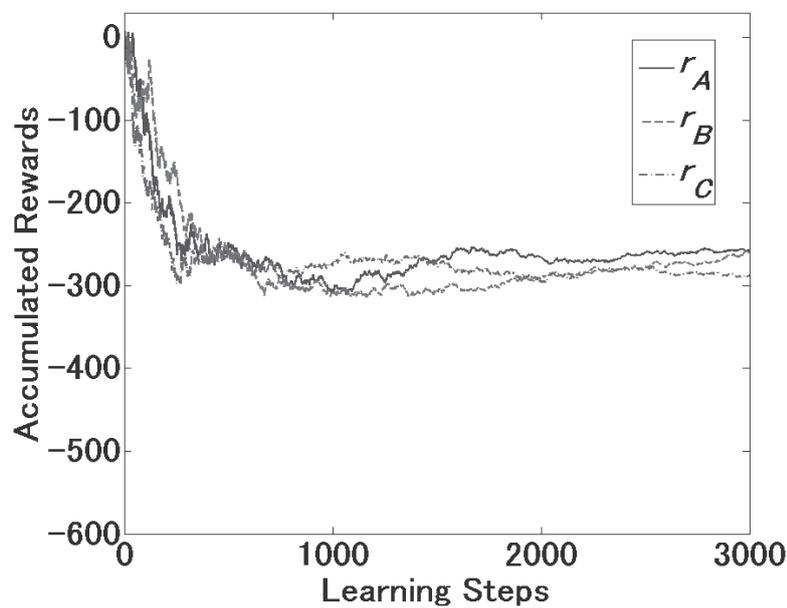


図5:各モジュールの累積報酬の推移

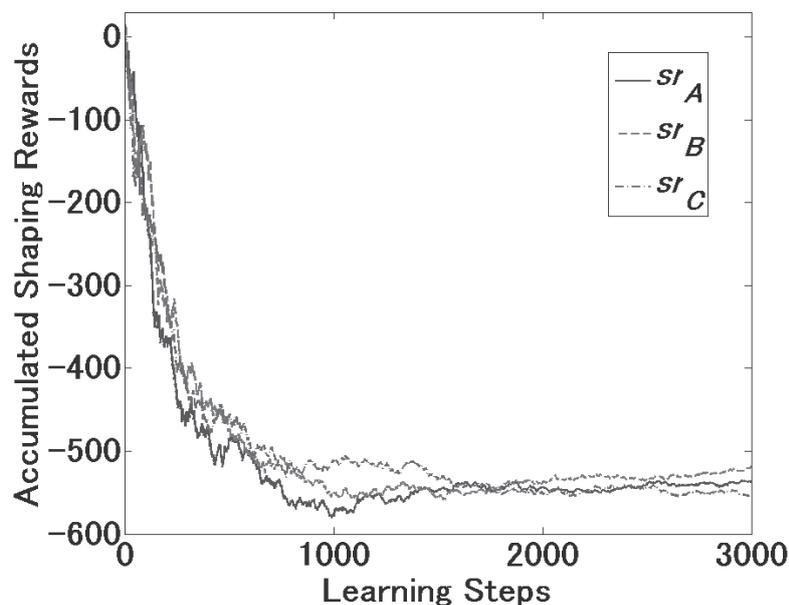


図6:各モジュールの累積シェイピング報酬の推移

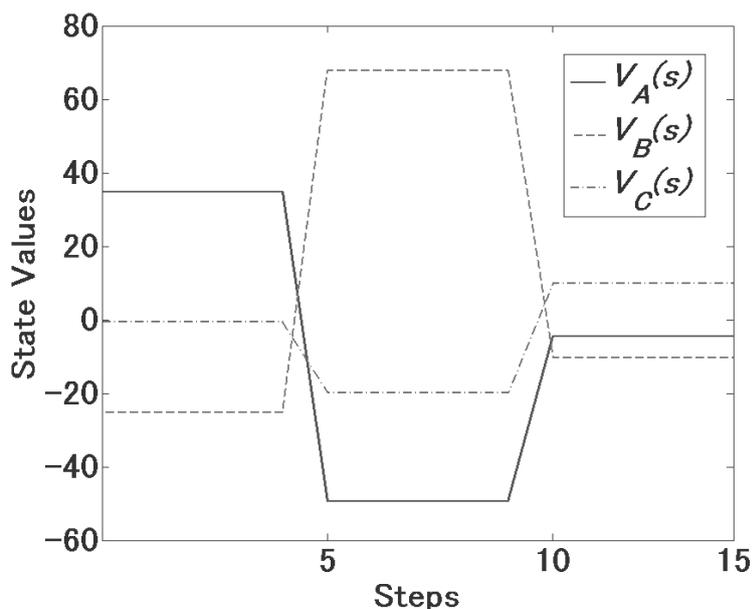


図7: 各モジュールの状態価値の推移

図7は、学習を終えた分人モデルに3人分の入力情報を与え、各モジュールからの出力である状態価値 $V(s)$ をプロットしたグラフである。Aさんの情報を与えたとき(ステップ数が1~5の間)、分人Aに相当するモジュールの出力が一番高く、分人Bと分人Cに相当するモジュールの出力は分人Aに比べて低くなっていることがわかる。また、Bさんの情報を与えた場合(ステップ数が6~10の間)およびCさんの情報を与えた場合(ステップ数が11~15の間)についても同様である。つまり、これは状態価値 $V(s)$ を観測することにより人を識別できることを示している。これらより、入力情報に対して複数ある分人の中から適切な分人が選択されていることが確認できる。

6. まとめ

本論文では、HRIの分野に分人という概念と行動分析学で使われるシェイピングの概念を導入し、モジュール型ニューラルネットワークとActor-Criticの手法を用いて、他者とのインタラクションを通して成長する分人モデルを構築した。そして、入力情報の定義と計算機シミュレーションにより、分人の形成と適切な分人の選択過程を確認した。

今後の課題としては、社会的な分人とグループ向けの分人の実装やインタラクションの定量化が挙げられる。また、実際のインタラクションを考慮した分人モデルの拡張をするべきである。そして最終的には、分人モデルの有効性を検証するために実機での実験が挙げられる。

近い将来、家庭にロボットが導入されたとき、我々は、家族一人一人とのインタラクションを通して成長し、学習し、家族一人一人に別々の対応ができることを望む。そのようなホームロボットの実現のためには、人間とロボットの円滑なインタラクションが可能になるモデルを開発しなければならない。分人を基盤としたヒューマン-ロボット・インタラクションにおけるモデルが開発されれば、ロボットが家族や医者といった人間社会の役割を担うことで、現在の人間とロボットの関係が大きく変わることもなるだろう。

参考文献

- [1] 井上 博允, 金出 武雄, 安西 祐一郎, 瀬名 秀明: “ロボット学創成”, 岩波書店, (2004)
- [2] 下山 勲, 柴田 智広, 生田 幸士, 波多 伸彦, 谷 淳, 松井 俊浩, 比留川 博久: “ロボットフロンティア”, 岩波書店, (2005)
- [3] 稲葉 雅幸, 加賀美 聡, 西脇 光一: “ロボットアナトミー”, 岩波書店, (2005)
- [4] 本田 幸夫: “ロボット革命 なぜグーグルとアマゾンが投資するのか”, 祥伝社, (2014)
- [5] T.Fong, I.Nourbakhsh, and K.Dautenhahn: “A Survey of Socially Interactive Robots,” *Robotics and Autonomous Systems*, Vol.42, pp.143-166 (2003)
- [6] M.A.Goodrich and A.C.Schultz: “Human-Robot Interaction: A Survey,” *Human-Computer Interaction*, Vol.1, No.3, pp.203-275 (2007)
- [7] 神田 崇行: “HRIにおけるソーシャルロボット研究の動向”, *日本ロボット学会誌*, Vol.29, No.1, pp.2-5 (2011)
- [8] 小野 哲雄, 今井 倫太, 江谷 為之, 中津 良平: “ヒューマンロボットインタラクションにおける関係性の創出”, *情報処理学会論文誌*, Vol.41, No.1, pp.158-166 (2000)
- [9] 小林 一樹, 山田 誠二: “行為に埋め込まれたコマンドによる人間とロボットの協調”, *人工知能学会論文誌*, Vol.21, No.1, pp.63-72 (2006)
- [10] 小嶋 秀樹: “ロボットの社会的発達と「心の理論」の獲得”, *情報処理学会研究報告*, ICS122-3, 13--18 (2000)
- [11] 柴田 崇徳: “アザラシ型ロボット・パロと人との相互作用に関する研究”, *日本ロボット学会誌*, Vol.29, No.1, pp.31-34 (2011)
- [12] 納谷 太, 篠沢 一彦, 小暮 潔: “人とロボットとの触覚インタラクション”, *情報処理*, Vol.44, No.12, pp.1227-1232 (2003)
- [13] 平野 啓一郎: “私とは何か「個人」から「分人」へ”, 講談社新書, (2012)
- [14] 鈴木 健: “なめらかな社会とその敵”, 勁草書房, (2013)
- [15] 中井 孝章: “分人主義宣言”, 日本教育研究センター, (2014)
- [16] M.Kawamura and K.Kobayashi: “An Action Selection Method Using Degree of Cooperation in a Multi-agent Reinforcement Learning System,” *Journal of Robotics, Networking and Artificial Life*, Vol.1, No.3, pp.231-236 (2014)
- [17] 奥田 健次: “メリットの法則 行動分析学・実践編”, 集英社新書, (2012)
- [18] 杉山 尚子: “行動分析学入門 ヒトの行動の思いがけない理由”, 集英社新書, (2010)
- [19] 杉山 尚子, 島宗 理, 佐藤 方哉, Richard W. Malott, Maria E. Malott: “行動分析学入門”, 産業図書, (1998)
- [20] T.Tanaka and K.Kobayashi: “Development of a Dividual Model Using a Modular Neural Network for Human-Robot Interaction,” *Journal of Robotics, Networking and Artificial Life*, Vol.2, No.1, pp.34-39 (2015)
- [21] R.Anand, K.Mehrotra, C.K.Mohan, and S.Ranka: “Efficient Classification for Multiclass Problems Using Modular Neural Network,” *IEEE Trans. Neural Networks*, Vol. 6, No. 1, pp.117-124 (1995)

- [22] R.S.Sutton and A.G.Barto (三上 貞芳, 皆川 雅章 共訳) : “強化学習”, 森北出版, (2000)
- [23] 前田 陽一郎, 花香 敏 : “調教の概念に基づいた自律移動ロボットの行動獲得支援手法”, 23rd Fuzzy System Symposium, pp.230-235 (2007)
- [24] 前田 陽一郎, 花香 敏 : “Shaping強化学習を用いた自律エージェントの行動獲得支援手法”, 日本知能ファジィ学会誌, Vol.21, No.5, pp.722-733 (2009)