

相互適応的 HRI を指向した学習型分人モデルの構築に関する研究

田中 利幸

指導教員：小林 邦和

1 はじめに

人間とロボットが対等なインタラクションをおこなうためには、ロボット自身が相手のことを考え、行動を変化させる様な適応的なインタラクションが重要となる [1, 2]. また、ロボットに相手がどのような人間なのかを学習させる仕組みが必要となる.

本研究では、文学者の平野啓一郎が提唱した分人という概念を HRI の分野に導入して、適応的なインタラクションを実現することを検討している. 分人は、環境や人間関係によって異なる「自分」になるという概念である. これにより、コミュニケーションを一層深めることができる. この概念を基に、先行研究では、機械学習の手法を用いることで、相手とのインタラクションを通して学習する分人モデルを構築している [3].

本稿では、先行研究で構築した分人モデルにおける問題点を解決するために、行動分析学の知見と行動決定レベルのアイデアを用いて拡張した分人モデルを示す. そして、計算機シミュレーションをおこない、拡張した分人モデルの有用性を述べる. 以降、第 2 節では、分人の概念について更に詳しく紹介する. 第 3 節では、分人モデルについて説明する. 第 4 節では、計算機シミュレーションとその結果の考察をおこなう. 第 5 節では、本研究のまとめと今後の課題について述べる.

2 分人

分人 (dividual) とは、対面する人間によって接し方が変化するという概念である [4]. また、分人は相手との反復的なコミュニケーションを通して自身の中に形成されていく人格であり、人間関係ごとに定義される. 人間は、この分人の概念を基に、人間関係によって変化する分人の集合体となる.

分人には、図 1 のように 3 つの種類があると考えられている. 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に所属している人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. 社会的な分人とグループ向けの分人を経て、最終的に生まれるのが、この特定の相手に向けた分人である. なお、人間は人間関係の数だけ分人を持っており、インタラクションの度にそれらを意識的ではなく無意識的に切り替えている. つまり、分人は中央集権型ではなく分散型のネットワークであると捉えることができる.

3 分人モデル

分人モデルは、分人の概念を基に、人間が持つ学習機構やインタラクション機構を数理的にモデル化したものである.

3.1 従来モデル

従来モデルは、枠組みとしてモジュール型ニューラルネットワーク (MNN) を採用している [5]. これを用いている理由は、分人の構成を表現するためである. 1 つのモジュールを 1 つの分人と捉えることにより、分人の概念における特徴的な分散型のネットワークを構築できると考えた. また、学習手法として強化学習の Actor-Critic 手法を採用している [6]. これを用いている理由は、分人が環境や人間関係の中で形成されることを表現するためである. 強化学習のメカニズムは、人間をはじめとする動物の脳内に存在することが明らかになってきており、分人の学習をモデル化していく上で妥当であると考えた [7, 8]. そして、想定される入力情報をカテゴリー要素と行動要素として定義し、分人モデルを構築した. しかし、当該モデルには、人間

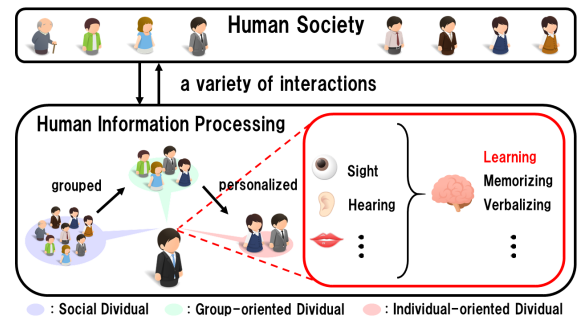


図 1 分人の概念図

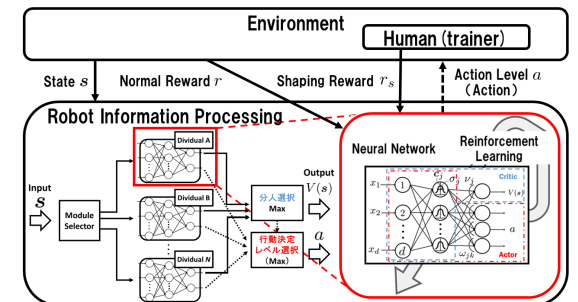


図 2 提案モデルの概念図

が教師となる系の強化学習が考慮できていない問題や行動決定の枠組みについて 3.2 節で述べる課題が残っており、検討の余地があった [3].

3.2 提案モデル

提案モデルは、従来モデルの拡張版である. 概念図を図 2 に示す. 各モジュールで分人を学習するために、提案モデルでは Actor-Critic 型強化学習と教師付き学習を適応する. そして、以下に示す拡張方針に沿って分人モデルの拡張をおこなう.

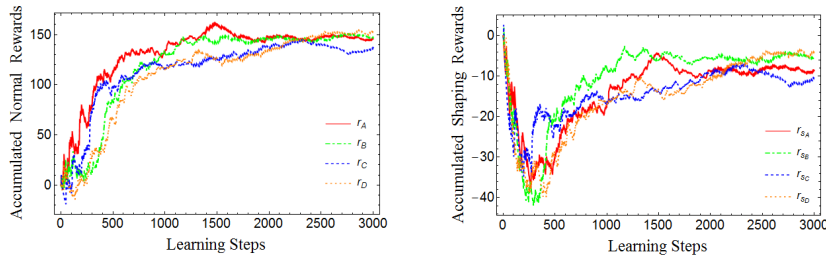
- 1. インタラクション的観点の導入：シェイピングの導入**
従来モデルにおいて、分人の概念のモデル化が不十分や学習的観点のみの評価しかできない、といった問題点があった. そこで、行動分析学におけるシェイピングという概念を導入し、分人モデルの拡張をおこなう [9, 10].
- 2. 段階的行動決定機構の構築：行動決定レベルの導入**
従来モデルにおいて、行動決定の枠組みに関して検討の余地があった. そこで、分人の種類に応じた行動決定を表現するために、行動決定レベルというアイデアを導入することで、行動決定機構の構築をおこなう [11, 12].

上記の拡張をおこない構築した分人モデル (提案モデル) の概念図を図 2 に示す. 分人モデルでは、式 (1) と式 (2) に示すように、critic に相当する分人識別部と actor に相当する行動決定部、それぞれの評価関数を定義する.

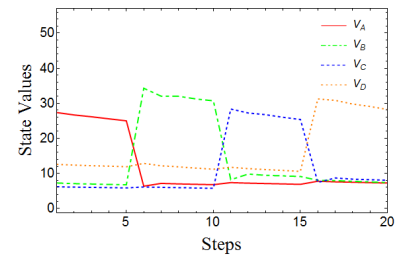
$$\delta_N = r + r_s + \gamma V(s') - V(s) \quad (1)$$

$$E_N = \frac{1}{2} \sum_{k=1}^3 (f(a_k) - t_k)^2 \quad (2)$$

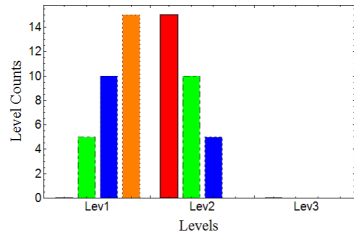
ここで、 r はノーマル報酬、 r_s はシェイピング報酬、 γ は割引率、 s は入力ベクトル、 $V(\cdot)$ は状態価値関数である. また、 a_k は行動値、 $f(\cdot)$ は活性化関数、 t_k は教師信号である. ここで、 N は分人の集合を表す. そして、これらの評価関数を最適とするように、学習パラメータを更新していく.



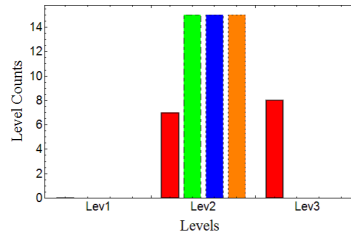
(a) 累積ノーマル報酬の推移 (b) 累積シェイピング報酬の推移
図3 事前学習における分人の学習



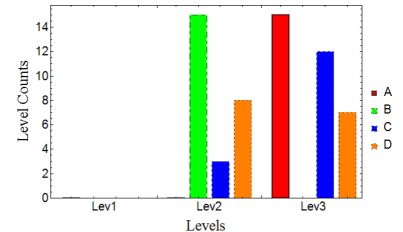
状態値の推移
図4 事後学習における分人選択



(a) インタクションの序盤



(b) インタクションの中盤



(c) インタクションの終盤

図5 事後学習における行動決定レベル選択

学習パラメータとなるのは、中間ユニットにおけるガウス関数の平均 c_j と出力層に紐づいている critic の重み ν_j と actor の重み ω_{jk} である。なお、式 (3) と式 (4)、式 (5) に示す更新式は勾配法に基づいて導出する。

$$c_j \leftarrow c_j + \zeta \delta \nu_j \frac{s - c_j}{\sigma_j^2} y_j \quad (3)$$

$$\nu_j \leftarrow \nu_j + \eta \delta y_j \quad (4)$$

$$\omega_{jk} \leftarrow \omega_{jk} - \rho (f(a_k) - t_k) f'(a_k) y_j \quad (5)$$

ここで、 ζ, η, ρ は学習係数、 σ_j^2 は分散、 $f'(\cdot)$ は活性化関数の導関数である。

4 計算機シミュレーション

前節にて構築した分人モデルの挙動を確認するために計算機シミュレーションをおこなう。ここで、分人の集合は $N = \{A, B, C, D\}$ とする。なお、本シミュレーションは、ロボットの適応的なインタクションの実現に向けた予備実験である。

4.1 シミュレーション設定

本研究では、日常生活での人間とロボットのインタクションを次のように想定する。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。このように、ロボットが対面している人間に関する情報を学習することで、ロボットは個々の対応を取ることができ、対等で適応的なインタクションがおこなえたと考える。それらを考慮し、本シミュレーションでは、1人あたり4パターン、すなわちカテゴリー情報は同じで行動情報が違う入力情報を4つ定義する。また、本シミュレーションで最も重要なのは、報酬の設定である。報酬の大きさは、各モジュールが適切に学習する見込み回数を考慮して設定する。また、報酬の与え方は、分人の概念と行動分析学における知見(人間の行動法則)に基づいて設定する。なお、報酬を含む各種パラメータ設定は、文献 [12] と同じである。

4.2 シミュレーション結果

図3の(a)、(b)から、累積報酬がある値に収束していることが確認できる。これより、各モジュールが強化と弱화를頻繁に繰り返して分人の学習をおこなっていることが示唆される。図4から、対面している人間に関する状態価値が、対面している人に関する状態価値が最も高い値を示すことが確認できる。これより、複数ある分人の中から適切な分人選択がおこなわれている

ことが示唆される。図5の(a)、(b)、(c)から、インタクションを重ねるにつれて、行動決定レベルがLevel 1からLevel 3に段階的に変化していることが確認できる。これより、適切な行動決定レベル選択がおこなわれていることが示唆される。

5 おわりに

本研究では、分人の概念と行動分析学の知見(人間の行動法則)、行動決定レベルのアイデアに基づき、モジュール型ニューラルネットワークとActor-Critic型シェイピング強化学習、教師付き学習を用いて分人モデルの再構築をおこなった。そして、予め設計した特定の人間の情報を入力情報とし、計算機シミュレーションをおこなった。シミュレーション結果より、各種累積報酬と状態価値、行動決定レベル累積値の推移を確認した。これらより、分人の学習と適切な分人選択、行動決定レベル選択がおこなわれることが示唆された。なお、これらは、HRIにおいてロボット自身が人間を分人として学習する仕組みを構築したことに相当する。

今後の課題としては、相手の意図を推定し行動決定をおこなうといった、実際のインタクションを考慮した分人モデルへの拡張があげられる。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をおこない、適応的なインタクションの実現を目指す。

参考文献

- [1] T. Fong, I. Nourbakhsh, and K. Dautenhahn: "A Survey of Socially Interactive Robots," Robotics and Autonomous Systems, Vol. 42, pp. 143-166 (2003)
- [2] 山田 誠二: 『人とロボットの〈間〉をデザインする』, 東京電機大学出版局, (2007)
- [3] 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)
- [4] 平野 啓一郎: 『私とは何か「個人」から「分人」へ』, 講談社新書, (2012)
- [5] 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)
- [6] 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)
- [7] M. Minsky, 安西 祐一郎: 『心の社会』, 産業図書, (1990)
- [8] 鮫島 和行, 銅谷 賢治: "強化学習と大脳基底核", パイオメカニズム学会誌, Vol.25, No.4, pp.167-171 (2001)
- [9] 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)
- [10] 田中 利幸, 鈴木 拓央, 小林 邦和: "HRIにおける分人モデルの構築-モジュール型ニューラルネットワークとシェイピング強化学習を用いた構成-", 第16回計測自動制御学会システムインテグレーション部門講演会予稿集, 3A3-2 (2015)
- [11] 田中 利幸, 鈴木 拓央, 小林 邦和: "HRIにおける段階的行動決定過程を考慮した分人モデルの構築", 第14回情報学ワークショップ (WiNF2016) ポスターセッション, A-02 (2016)
- [12] 田中 利幸, 鈴木 拓央, 小林 邦和: "HRIにおける段階的行動決定機構を備えた分人モデルの構築", 電気学会システム研究会, No.ST-16-052, pp.59-64 (2016)