

複数人直接教示によるロボットの適応的行動スキルの抽出とインタラクション解析

Extraction and Analysis of Adaptive Action Skills of Robot under Social HAI

片上大輔*¹ 生和太郎*¹ 新田克己*¹
Daisuke Katagami Tarou Nyuwa Katsumi Nitta

*¹東京工業大学大学院総合理工学研究科

Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

In this paper, we paid attention to Social Adaptation Problem in Human-Agent Interaction and defined Social HAI model. We introduced sequential clustering method to our conventional method and showed the highly accurate cluster as the results by social interaction with multiple instructors. Moreover, we confirmed social mutual adaptation by analyzing of interaction between subjects and a robot implemented our method.

1. はじめに

人間は高度な適応能力をもち、コンピュータ相手であっても、ロボットでも、それが道具であっても適応すると言われている。一方、擬人化エージェントやソフトウェアエージェント、ロボットに代表されるエージェントも人間に自然に適応するための研究・開発が活発に行なわれている。エージェントは家庭や職場など人間社会で活躍はじめており、今後多くの人間に囲まれて生活空間を共にしていくことが予想される。この人間とエージェント間のインタラクション設計問題はHAI (Human-Agent Interaction) と呼ばれ、近年特に日本を中心に活発に研究が行なわれている。

しかし、これまでは主に人間という個別のユーザを対象としており、社会に存在する複数の多様性の高い人間に焦点をあててはいなかった。前述の人間-エージェント共生社会を想定すると、HAIを考える上で社会的環境を考慮することが重要である。

本研究では、HAIの社会的モデルを考え、複数人とインタラクションを行なう上で問題となる社会的HAI適応問題を取り上げる。本稿では、この適応設計の例として、複数人教示における行動スキルの獲得と社会的HAIのインタラクションの解析を行なう。

2. 社会的環境におけるHAI

2.1 適応のためのHAIモデル

山田らはHAIにおける適応のためのインタラクション設計について、人間-エージェント間でのインタラクションを主に構成要素としてモデル化を行なっている[山田 03]。ここでは、モデルの構成要素は、人間の行動 B_i^H 、エージェントの行動 B_i^A 、行動モデル $M(B)$ 、インタラクション I_i で示される。HAIにおいては、人間は対象であるエージェントに自然に適応が生じる。また、エージェントがユーザとの自然なインタラクションを実現するには、ユーザへの個人適応、つまりユーザモデル $M(B_{i-1}^H)$ を学習して、それに基づいて自身の行動を修正する必要がある。このように人間とエージェントで互いに相手に適応する状況を相互適応 (mutual adaptation) と呼ぶ。

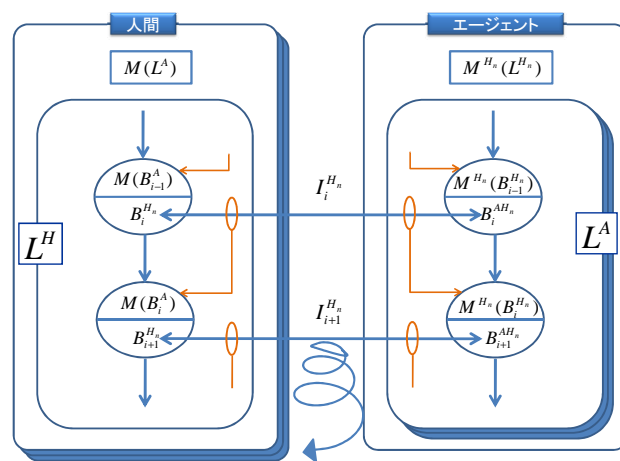


図 1: 社会相互適応モデル ([山田 03] を修正)

ここで、相互適応を促進するインタラクション設計の対象は、エージェントの行動 B_A 、エージェントの学習 L_A である。このモデルにおいては、人間の適応の促進、エージェントの適応の促進、相互適応の促進の3つのインタラクション設計が目的とされる。

2.2 社会的HAIモデル

上記のようなインタラクション空間に人間が複数存在する社会的環境を考えると、前述のHAIモデルは図1のように表現できる。社会的環境におけるHAI (Social-HAI) モデルの構成要素は、人間の行動 $B_i^{H_n}$ 、エージェントの行動 $B_i^{AH_n}$ 、行動モデル $M(B_i^{H_n})$ 、インタラクション $I_i^{H_n}$ で示される。S-HAIにおけるインタラクション設計が従来のHAIにおける設計と大きく違うところは、人間からのインタラクション $B_i^{H_n}$ の方針が人間によって異なることである。このインタラクションの方針の違いを、エージェントの学習に用いるために何らかの形で統合を行なう必要がある。山田らは、これをHAIにおける学習の汎用性と呼び、これによりインタラクションの促進することができるとしているが、同時にこのインタラクション設計に関してはほとんど研究例がないとも指摘している[山田 03]。

連絡先: 片上大輔, 東京工業大学大学院総合理工学研究科, 神奈川県横浜市緑区長津田町 4259 J2-53, 045-924-5218, katagami@ntt.dis.titech.ac.jp

一方、エージェントから人間へのインタラクション $B_i^{AH_n}$ を考えると、複数の人間にどのように適応するかなどの問題も生じる。エージェントは、社会的環境においては単に個人適応をすればよいわけではなく、エージェントから人間へのインタラクションに対する様々なタイプの人間の適応による相互適応の影響を考慮に入れ繰り返されるインタラクションを設計しなければならない。このように、多様にかつカスケードに積み重なる適応のためのインタラクションにより、複数の人間とエージェントの学習 L^H と L^A がスパイラルに促進するような適応設計が、S-HAI のインタラクション設計として重要な課題であるといえる。

2.3 社会的 HAI における適応問題

複数の教示を統合する研究にスキルサイエンスとして知られる研究の分野がある。これは、複雑な動作をロボットに学習させるために、複数回に渡り教示動作を与え重要な部分を抽出し汎化することを目的としている。Ogawara らはマニピュレータロボットに、人間からの複数回の直接教示により目的のタスクを達成する動作を汎化し獲得させている [Ogawara 03]。この研究ではタスクを自動的にサブタスクへと分割し、DP マatching法により冗長な要素を省き必須要素のみを抽出している。しかし、教示は全て同じ方針で目的を達成すると設定しており、人間の行動モデル $M(B_i^{H_n})$ がすべて同じと仮定していると考えることができる。

我々はこれまでに、固有の行動モデル $M(B_i^{H_n})$ をもつ複数人の教示者からの教示動作が与えられた際にロボット自身の身体性に基づき最適な方策の選択を行なうシステムの構築を行ってきた [Kotake 07]。しかし、教示データが類似していても動作の評価が類似していない場合があった。また、教示データを集めてのオフラインシステムであり、対話的なインタラクションによる適応はできなかった。これにより、S-HAI 環境においては、逐次的に行なわれよう社会的相互適応が実現できていなかったといえる。そこで、本稿では複数の評価値を用いた固有行動モデルの汎用化の実現と逐次的クラスタリングとフィードバックによる対話的なインタラクションを導入し、より適切な社会的 HAI 適応を行なうことを目的とする。

3. ロボットのための社会的適応システム

従来システム [Kotake 07] に対して、本研究のシステムは次の二点を特徴とする。表示モジュールと評価値算出モジュールを含むロボットシミュレータを追加することで、教示時の動作の再現、教示データへの評価値の付与、教示後の代表教示の表示による教示者へのフィードバックを行うことができる。また、教示者へのフィードバックを行なうため、教示者が教示をする毎に代表教示データを作成する。そのために、DP マatching, クラスタリング, 平均化モジュールを統合し、逐次クラスタリングモジュールを構築した。

3.1 システムの概要

提案するシステムの概略図を図2に示し、以下に説明を行う。

- 1) 教示者が、ロボットに対し直接手を使ってロボットを動かして教示を行う。
- 2) ロボットシミュレータにおいて、表示モジュールが画面上にロボットの教示時の行動の表示を出力し、評価値算出モジュールがシミュレータ上のセンサ情報を取得し、各教示データに対する評価値を出力する。
- 3) 教示者が画面に出力された自分の教示動作と評価値に対して評価を与える。

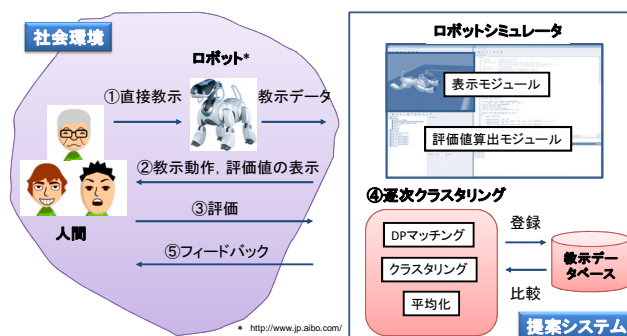


図 2: システム概要

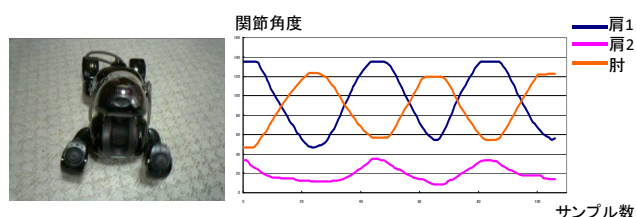


図 3: 3 自由度の腕関節に与えられる直接教示データ

4) 逐次クラスタリングモジュールにおいて、DP マatchingモジュールが現在の教示データとデータベースにある過去の複数の教示データを入力とし、DP マatching法により、2つの時系列データの波形の距離を求め出力する。クラスタリングモジュールでは、波形距離と、各教示データの評価値を入力とし、各教示データ間の距離を算出する。また、クラスタリング手法を用いて距離から各教示データをクラスタ化する。平均化モジュールは、クラスタリングモジュールによって結合された教示データを入力とし、各クラスタ内の教示データを平均化し、代表教示データとして出力する。

5) 逐次クラスタリングモジュールで出力されたクラスタの代表教示データの中から、現在の教示データが含まれるクラスタの代表教示データを表示モジュールに入力し、ロボットシミュレータ上に表示することにより、教示者へのフィードバックを行う。

3.2 教示データ

教示データはロボットの動作を扱うために、複数次元の時系列データからなる関節角度のサンプリングデータを用いる。図3に教示データの一例として、3自由度の関節を持つロボットの教示データを示す。横軸はサンプル数を表し、縦軸は関節角度を表す。この教示データをロボットシミュレータの評価値算出モジュール、表示モジュール、逐次クラスタリングモジュールの DP マatchingモジュールへ入力する。

3.3 ロボットシミュレータ

ロボットの状態の把握やユーザへのフィードバックに利用するロボットシミュレータとして cyberbotics 社が提供する、Webots を利用する。対象とするロボット、環境などを VRML (Virtual Reality Modeling Language) を用いてシミュレータ上に実現し、その動作をプログラミングすることで、環境での移動ロボットの動作を検証するのに利用することができる。

また、重力、摩擦係数などの物理情報を設定し、力学の法則により、速度、加速度、力などを計算することができる。また、あらかじめ khepera や AIBO など様々なロボットがプロジェクトとして組み込まれている。我々は、このシミュレータ上に表示モジュールと評価値算出モジュールの2つの拡張モジュールを追加した。

3.3.1 表示モジュール

表示モジュールは、教示データが入力された際に、ロボットシミュレータ上のロボットにその時系列データを入力して再現し、出力する。教示者がロボットに対して教示を行っている際には、画面上に教示時のロボットの動きを再現する。これにより、教示時と実行時の動作のギャップを抑えることができる。また、教示後には教示データの入ったクラスタの代表教示データを表示することにより、教示者へのフィードバックを行う。

3.3.2 評価値算出モジュール

評価値算出モジュールは、教示データを入力し、シミュレータ上で得られるセンサ情報からその教示行動に対する評価値を算出し、出力する。利用できる Webots のセンサとしては、タッチセンサ、距離センサ、カメラ、レーザセンサ、GPS などがある。評価値の例としては、移動するロボットであれば、移動距離、移動時間、速度、加速度などが実現できる。

3.4 逐次クラスタリングモジュール

逐次クラスタリングモジュールは、DP マッチングモジュール、クラスタリングモジュール、平均化モジュールからなる。

3.4.1 DP マッチングモジュール

DP マッチングモジュールは、現在の教示データ、過去の教示データを入力とし、各教示データの波形の距離を出力する。

DP マッチング法は時系列データに対し、データ間の距離を求めるための計算法である [Sakoe 78]。二つの時系列データのうち、一つのデータの経路1点をもう一つのデータの複数の経路点と対応付けを行うことで、時間方向に非線形な伸縮をすることが可能となっている。これにより、サンプリング数の違う時系列データ同士の距離を求めることが可能である。

サンプリング数がそれぞれ I, J の教示データ $A = a_1, \dots, a_I, B = b_1, \dots, b_J$ 間の距離 $D(A, B)$ を求める場合を考える。 $a_i, b_j (1 \leq i \leq I, 1 \leq j \leq J)$ は、センサから得られる測定値やその測定値から計算される特徴の値などからなる多次元ベクトルである。各 i と $j (1 \leq i \leq I, 1 \leq j \leq J)$ に関する e_{ij} からなる行列 E を (1) 式で定義する。ここで、教示データ A, B 中のサンプリング点 a_i と b_j の距離 e_{ij} を (2) 式とする。

$$E = \begin{pmatrix} e_{11} & \cdots & e_{1I} \\ \vdots & \ddots & \vdots \\ e_{1J} & \cdots & e_{IJ} \end{pmatrix} \quad (1)$$

$$e_{ij} = |a_i - b_j| \quad (2)$$

e_{11} から e_{IJ} まで、 $e_{(i+1)j}, e_{i(j+1)}, e_{(i+1)(j+1)}$ のうち最も小さい距離を足していき、 e_{11} から e_{IJ} までの最短の経路長 $D_{min}(A, B)$ を (3) 式のように求める。

$$D_{min}(A, B) = \sum_{i=1}^I \sum_{j=1}^J e_{ij} + \min(e_{(i+1)j}, e_{i(j+1)}, e_{(i+1)(j+1)}) \quad (3)$$

教示データ A, B 間の距離 $D(A, B)$ は、 e_{11} から e_{IJ} までの最短の経路長 $D_{min}(A, B)$ を経路点の総数 K で割った (4)

式によって求める。

$$D(A, B) = \frac{D_{min}(A, B)}{K} \quad (4)$$

求められた教示データ A, B 間の波形の距離 $D(A, B)$ を DP マッチングモジュールの出力とする。

3.4.2 クラスタリングモジュール

クラスタリングモジュールでは、教示データと過去の教示データを入力とし、各クラスタにデータを振り分ける。複数人による教示の方針の数があらかじめ決まらないため、クラスタリング手法には階層的クラスタリングを採用した。

従来手法ではクラスタリングの際に用いる教示データ間の距離として、各教示データの時系列の波形の距離を用いていた。ここで、教示データ A, B 間の距離は $S(A, B)$ 、波形の距離は $D(A, B)$ とすると $S(A, B) = D(A, B)$ で表される。

本手法では、行動の評価値を含めた距離を定義することで、教示動作の動きと共に動作の評価値を考慮したクラスタリングを行う。動作の評価値は教示行動の目的に対し複数考えられるが、例として教示データの評価値が2つあった場合を考える。波形の距離と衝撃値という異なる次元の距離を用いるため、適合度関数としてよく用いられる重み付け線形和を用いた (5) 式)。ここで、距離は $S(A, B)$ 、波形の距離 $D(A, B)$ の正規化した値を $D'(A, B)$ 、教示データ間の評価値1の差を正規化した値を $E'_1(A, B)$ 、評価値2の差を正規化した値を $E'_2(A, B)$ 、各評価値に付与する重みを α, β とおく。

$$S(A, B) = D'(A, B) + \alpha E'_1(A, B) + \beta E'_2(A, B) \quad (5)$$

3.4.3 平均化モジュール

クラスタリングモジュールにおいてクラスタに格納された教示データは同じ方針で教示をもつと考えられる。しかし、同じ方針といえども個人差があるために多少の差や個人の癖があることが予想され、その影響を取り除く必要がある。また、教示数圧縮という観点においても、各クラスタから代表される1つの教示データを出力する必要がある。平均化モジュールでは、各クラスタごとに教示動作データを平均化することにより、クラスタを代表する代表教示データを出力する。

一般に時系列データの場合、単純に時間に沿って平均化すると時間軸のずれなどにより人間の感覚に反する平均化をしてしまうことが多い。Nakamoto らはこの問題に対し、DP マッチング法により時系列データの平均化を行っている [Nakamoto 02]。本研究ではこの手法を用いて教示動作データの動的時間伸縮を行い平均化を行う。

4. 実験

本研究では、S-HAI における2つの実験を行った。実験1では、行動モデルの汎用化として、動作の評価値を含めたクラスタリングにより、精度の高いクラスタの生成が可能であることを示す。実験2では、逐次クラスタリングによって対話的な社会相互適応が実現できていることを示す。

4.1 実験設定

タスクは、人間と異なる身体性 (20 自由度) を有する大型ペットロボット AIBO(ERS-7) を用いる。本実験では、AIBO の前足 (各3自由度) を直接教示 (AIBO の前足を直接掴み動かす教示方法) により、可能な限り前進させることを目的とする。タスクの評価値は、ロボットの前足距離 E_1 と衝撃値 E_2 とし、その重みはそれぞれ $\alpha, \beta = 1 - \alpha$ とした。サンプリング間隔は 128ms で実験を行なった。

表 1: 重み α を変えた際の最大クラスタにおける評価値

	n	前進距離	分散	衝撃値	分散
従来手法	9	0.141	2.21E-03	1.349	0.482
$\alpha = 0.0$	6	0.169	8.44E-04	0.356	0.005
$\alpha = 0.3$	5	0.106	8.53E-04	0.489	0.061
$\alpha = 0.5$	5	0.183	7.01E-04	2.102	0.213
$\alpha = 0.8$	5	0.168	7.01E-04	1.325	0.213
$\alpha = 1.0$	7	0.207	5.20E-04	1.018	0.334

4.2 実験 1: 行動モデルの汎用化

事前に行なった同タスクの 37 人の被験者のオフライン教示データをロボットシミュレータ上で再現し、前進距離、衝撃値を行動の評価値として各教示データに与え、疑似 F 値統計量を基準にクラスタリングを行った。教示データ間の距離計算に用いる評価値の重みを 0.0~1.0 まで変化させ、クラスタリングを行った。また最大クラスタの評価値の分散を従来手法と比較した。また、各重みでできたデンドログラムの中で最大のクラスタの代表教示データを作成した。代表教示データをロボットシミュレータに入力し、各代表教示データに評価値を与えることで各重みでの代表教示データの評価値を従来手法と比較した。表 1 に従来手法と本手法で距離計算の際に用いた重みを 0.0, 0.3, 0.5, 0.8, 1.0 に変化させた際に得られた最大クラスタ内の教示データのデータ数 n , 代表教示データの前進距離と衝撃値の評価値, 各評価値の分散を示す。クラスタ内の各評価値の分散が、前進距離 ($\alpha = 1.0$) で 4 分の 1, 衝撃値 ($\alpha = 0.0$) で 100 分の 1 にまで改善されており、評価値に応じたクラスタが構築されていることがわかる。

4.3 実験 2: 社会的相互適応

大学院生 4 人を対象としてインタラクションの解析をおこなった。実験手順を以下に示す。

1) 被験者が AIBO に対し、前足のみの前進運動を直接教示する。

2) システムは得られた教示データと過去 37 人の教示データから逐次クラスタリングを行い、ロボットシミュレータ上で AIBO の代表教示動作を出力する。

3) 被験者はシミュレータ上に出力された AIBO の動作を確認し 1 に戻る。教示を規定回数行ったら終了する。

教示と出力の繰り返しにより改善された前進距離の典型的な 2 被験者のインタラクションの例を表 2 に示す。被験者 1 は 2 回目と 3 回目の教示回数では前進距離の値が教示動作に比べ、出力動作の方が小さい。しかし、3 回目の出力動作の前進距離の大きさに比べ、4 回目の出力動作の前進距離の値が大きい。アンケートでは、「出力動作が自分の考える教示動作と同じであり良い動作である」と回答した。初めは出力された動作が被験者 1 の考える動作と異なるため、自分の考える動作を入力し続けたが、4 回目に自分の考える動作が AIBO の出力動作として現れたと考えられる。

被験者 2 は、1 回目の教示動作に比べ、出力動作の前進距離の値が大きい。2 回目では、1 回目に比べ教示動作、出力動作ともに前進距離が大幅に大きい。アンケートでは「出力動作の方が地面に肘が触れないように腕を回していた」と回答した。3 回目以降も大きな前進距離を維持している 2 回目の教示で教示動作、出力動作ともに前進距離が大幅に大きくなっているのは、出力動作を参考にすることで教示動作の改善を行ったと考

表 2: 被験者とロボットの相互適応による前進距離の改善

回数	被験者 1		被験者 2	
	教示動作	出力動作	教示動作	出力動作
1	9.2	11.8	11.8	16.0
2	13.3	11.5	23.4	22.5
3	15.2	10.2	19.4	19.3
4	15.4	16.4	22.0	22.6

えられる。

教示動作とロボットが出力した動作が異なる場合教示者がその教示方針を持続することでその教示者独自のクラスタを作成し、新たな方針が教示者とロボットの間で創発される。一方、教示者が出力動作（代表教示）に適応し、教示方針を柔軟に変更することで、当初の教示方針よりもよい動作が得られることもある。どちらも教示者が、教示動作とロボットの出力動作の違いを認識し、相互にインタラクションすることで、よりよい出力動作を対話的に獲得していると言える。

5. おわりに

本稿では、HAI の社会的モデルを定義し、複数人教示における行動スキルの獲得とインタラクションの解析を試みた。ロボットの行動スキルに評価値と逐次のクラスタリングを導入し、行動モデルの汎用化として複数ユーザとの対話的な教示により精度の高いクラスタが生成される事を示した。また、社会的相互適応を被験者とロボットのインタラクション解析により確認した。今後は社会的相互適応のスパイラル促進を目指した適応設計を行なう。

参考文献

- [Ogawara 03] K. Ogawara, et. al.: Extraction of Essential Interactions Through Multiple Observations of Human Demonstrations, IEEE Trans. on Industrial Electronics, Vol.50, No.4, pp.667-675 (2003)
- [Kotake 07] M. Kotake, D. Katagami, K. Nitta: Acquisition of behavioral patterns depends on self-embodiment based on robot learning from multiple instructors, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.11, No.8, pp.989-997 (2007)
- [Sakoe 78] H. Sakoe, S. Chiba: Dynamic Programming Algorithm Optimization for Spoken Word Recognition, IEEE Transaction on Acoustics, Speech, and Signal Processing, Vol.ASSP-26, No.1, pp.43-49 (1978)
- [山田 03] 山田 誠二, 角所 考: IDEA: 適応のためのインタラクション設計, 日本知能情報ファジィ学会誌, Vol.17, No.3, pp.279-288 (2005)
- [Nakamoto 02] K. Nakamoto, E. Suzuki: Fast Clustering for Time-series Data Based on a TWS Tree, Proc. 48th SIG-FAI, Japanese Society for Artificial Intelligence, pp.9-14 (2002)