

人間とエージェントの相互適応

Mutual Adaptation between a Human and an Agent

山田誠二*1
YAMADA Seiji

山口智浩*2
YAMAGUCHI Tomohiro

*1 国立情報学研究所
National Institute of Informatics

*2 奈良工業高等専門学校 情報工学科
Nara National College of Technology

This paper describes how to design interaction in human-agent mutual adaptation. In human-agent interaction, a user and an agent mutually acquire adaptive behavior to the other. We propose important properties to design the interaction on parallel adaptation and adaptation interferer, and a test bed framework in which a human and an life-like agent acquire the other's mind mapping through a mutual mind reading game.

1. 人間とエージェントの相互適応：HAMA

擬人化エージェント、ペットロボットなどの普及により、人間とエージェントが接する機会が増えているおり、このような人間とエージェントのインタラクション HAI(Human Agent Interaction) について様々な研究が始まっている [山田 02]。HAI において、人間は不可避的にエージェントに適応を試み、反対にエージェントは人間に適応してパーソナライズする機能が組み込まれることで、人間とエージェントが互いに適応する状況が生じる。我々は、これを人間とエージェントの相互適応 HAMA(Human Agent Mutual Adaptation) と呼び、それを促進するためのインタラクション設計の研究を始めている [Yamada 02]。HAMA の促進により、人間とエージェントは自然にかつ効率よく相手の状態を理解できるようになり、相手の気持ち考えた円滑な付き合い方が実現される。また、HAMA が効果的に実現されるには、人間とエージェント間の情報の表現、伝達方法、また、エージェントの行為選択、適応(学習)アルゴリズムの特性などが重要でありが、本研究では、このような要素の設計を広くインタラクション設計と呼ぶ。

以上のような考え方にに基づき、HAMA のためのインタラクション設計において、以下の性質を実験的に調査することが本研究の目的である。

1. 同時並列性: HAMA では、相互適応を直列的にも実現できるが、同時並列的に実現することで、様々なヒューリスティクスを利用でき、質と効率を向上できる可能性がある。
2. 適応干渉の許容: 自分の適応(学習)対象が相手に対しても操作可能である場合、自分の学習結果が相手の学習を妨害してしまう適応干渉が生じる。この適応干渉は、マルチエージェント学習では回避される傾向にあるが、HAMA の場合は、人間の適応能力がエージェントを遥かに凌駕しており、ある種の適応干渉を許容することで、相互適応を促進できる可能性がある。

2. 人間と擬人化エージェント間のマインドマッピングの相互適応

本節では、本研究の実験で用いる HAMA の枠組みである、人間と擬人化エージェント間のマインドマッピングの相互適

応について説明する。近年、活発に研究開発されている擬人化エージェントと人間がインタラクションをもつとき、人間はエージェントの表情などを基にエージェントの気持ち、負荷など内部状態(マインドと呼ぶ)を推定するようになる。反対に、エージェントが人間の表情などから人間のマインドを推定しようとする。このように人間と擬人化エージェント間で HAMA が生じる。

しかし、人間がエージェントの表情(図1)を理解することも、逆にエージェントが人間の表情を理解することも一般には容易ではない。この表情からマインドへの写像をマインドマッピングと呼ぶ。本研究では、人間と擬人化エージェントのマインドマッピングの相互適応を HAMA のテストベッドとして、前述の2つの性質の評価実験を行う。

2.1 相互読心ゲーム

同時並列性を満たすために、相手の表情からマインドを推定することを交互に行う相互読心ゲームを提案する。これは、以下のような手続きを繰り返す協調ゲームであり、人間とエージェントのペアの間で競争する。人間と擬人化エージェントの双方が、相手の異なるマインドについて連続して一定回数以上正しく言い当てれば、ゲームは終了する。

1. 相手の表出した表情から相手のマインドを推定し、推定したマインド名を相手に提示する。
2. 相手の推定したマインド名に対し、“正しい”、“間違い”を相手に伝える。エージェントの場合は、さらに正しいマインド名を人間に提示。

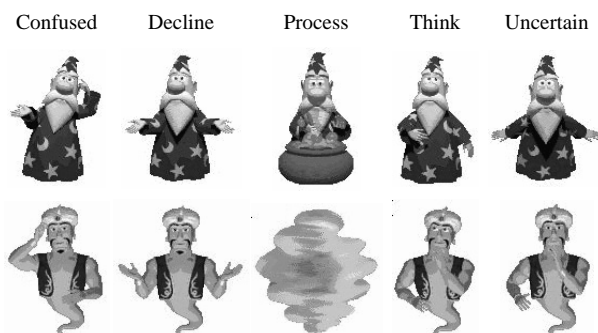


図1 MS エージェントの表情

連絡先: 山田誠二, 国立情報学研究所, 〒101-8430 東京都千代田区一ツ橋 2-1-2, seiji@nii.ac.jp

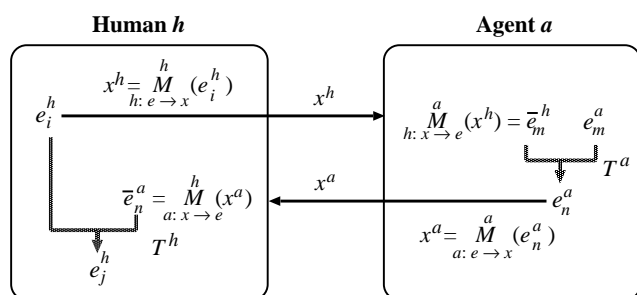


図 2 エージェントと人間のインタラクション

2.2 基本概念

相互読心ゲームを構成する基本概念を以下に示す．詳細は，[Yamada 02] を参照．

- マインド状態 s_h, s_a : 人間 (エージェント) のマインドの状態を表す変数．
- 基本マインド $E^h = \{e_1^h, \dots\}, E^a = \{e_1^a, \dots\}$
- 基本表情 $X^h = \{x_1^h, \dots\}, X^a = \{x_1^a, \dots\}$: x は，それぞれの基本マインドから表出される基本表情の集合．
- マインドマッピング $M_{h:x \rightarrow e} = \{x_i^h \rightarrow e_j^h, \dots\}, M_{a:x \rightarrow e} = \{x_i^a \rightarrow e_j^a, \dots\}$: 基本表情から基本マインドへの 1 対 1 写像．
- 表情マッピング $M_{h:e \rightarrow x}, M_{a:e \rightarrow x}$: 基本マインドから基本表情への写像で，マインドマッピングの逆写像．
- マインド遷移関数 $T^a(c), T^h(c)$: 次のマインドを出力する文脈 c の関数である．文脈 c は，現在のマインド，過去のマインド履歴，成功率などをパラメータとする．

以上の定義により，人間とエージェント間の表情によるインタラクションの枠組みは，図 2 のようなる．エージェントの学習対象は，人間のマインドマッピング $M_{h:x \rightarrow e}$ と自分の表情マッピング $M_{a:e \rightarrow x}$ の両方である．このエージェントの表情マッピングは，人間が学習するエージェントのマインドマッピングの逆写像であり，ここで適応干渉が生じる．また，人間側の学習には制約を与えないため，エージェントのマインドマッピングと自分の表情マッピングの両方を学習することで，人間の表情マッピングについて適合干渉が生じる．

3. エージェントの学習

現在観測される相手の表情に対する相手のマインドが特定できれば，以下の事例ベース学習により，エージェントは，人間のマインドマッピングを学習できる．

1. マインド状態 e を同定して，そのときの表情 x を CCD カメラで獲得し， $e \Rightarrow x$ をサンプリングする．
2. 集められた $\{e \Rightarrow x, \dots\}$ を事例として，NN (Nearest Neighbor) 法で， x から e を推定する．

次に，エージェントは，以下のように自分の表情マッピングを修正して，人間に適応することができる．

1. エージェントの各基本マインドについて，初期表情を設定，間違い回数を 0 にリセット．



図 3 実験環境

2. 間違い回数が N_E になると，最も間違ったマインド e_E とその表情 x_E の表情マッピングを生成．元々 x_E にマッピングされていた基本マインドは，別の基本表情にマッピングする．間違い回数を 0 にリセット．

一方，人間の学習に対して制約を与えていない．その結果，人間は，エージェントのマインドマッピングと自身の表情マッピングの両方を学習する傾向がある．

4. 実験

実験環境は，図 3 のようなプロトタイプが完成している．ユーザはモニタ上に現れるエージェントを見て，エージェントの方は人間を CCD カメラでモニタすることができる．相互読心ゲームを行なうことにより，擬人化エージェントと人間のマインドマッピングが相互に学習され，相互適応が実現される．

今後は，同時並列性の効果を調べるために，相互読心ゲームと直列的な相互適応との比較を行い，さらに，適応干渉の効果を調べるために，3. の間違い回数の閾値 N_E を変化させた場合に相互適応がどのようになるかを実験的に調べる．これらの実験において，学習効率と認知的負荷を評価基準として用いる．

5. まとめ

本稿では，HAMA におけるインタラクション設計における重要な特性と，それらの効果を調べるためのテストベッドである人間と擬人化エージェントのマインドマッピングの相互適応の枠組みを提案した．今後は，本枠組みをもとに，実験的な検証を行う．

参考文献

- [山田 02] 山田, 角所, 新田: 特集: HAI ヒューマンエージェントインタラクション, 人工知能学会誌, Vol. 17, No. 6 (2002).
- [Yamada 02] Yamada, S. and Yamaguchi, T.: Mutual Learning of Mind Reading between a Human and a Life-like Agent, in *The Fifth Pacific Rim International Workshop on Multi-Agents*, pp. 138-150 (2002).