


 原著論文

人間と擬人化エージェントによる マインドマッピングの相互適応[†]

山田 誠二^{*1}・山口 智浩^{*2}

擬人化エージェントと人間との間に効果的なコミュニケーションを実現するためには、双方が相手の表情から相手の感情、負荷状態などの内部状態(マインド)を同定し、相手のマインドに沿った行動をとることが重要である。しかし、人間と擬人化エージェントが互いに相手の表情からマインドを同定することは、一般に困難である。なぜならユーザはエージェントの表情の解釈において、個人差、文化的差異などをもつため、すべてのユーザにとってマインドを容易に推定できる表情をもったエージェントを設計することは難しいからである。

本論文では、人間と擬人化エージェントがマインドマッピングをお互いに学習して相互適応を実現する枠組みを提案する。マインドマッピングとは、人間や擬人化エージェントの表情からマインドへの写像である。人間とエージェントは、相互読心ゲームと呼ばれるゲームを行うことにより、相手のマインドマッピングを学習していき、相互適応が実現される。この相互適応において、エージェントは、観測された人間の顔画像を事例とした事例に基づく学習により、マインドマッピングを学習し人間に適応する。一方、人間側は自由に学習をしてエージェントに適応する。PCとCCDカメラを用いてこの枠組みを実装し、我々の枠組みで実際にマインドマッピングの相互適応が実現されることを実験的に示す。また、マインドマッピングの相互適応を高速化するヒューリスティックスを提案する。

キーワード：相互適応，擬人化エージェント，ヒューマンエージェントインタラクション

1. はじめに

近年、マイクロソフトエージェント[9]、インタフェースエージェント[8]に代表されるような擬人化エージェント(life-like agent)の研究、実用化が盛んに行われている[5]。典型的な擬人化エージェントは、Webのオンラインショッピングページ上に現れ、品物、個数などの入力の支援をしたり、オフィスアプリケーション上に現れて、その操作方法の説明をする。

このような様々な擬人化エージェントの研究を通じて、エージェントの感情や処理の負荷などの内部状態を組み込み、さらにそれらを人間にうまく伝達することが、人間とのインタラクションにおいて重要な働きをすることがわかってきた[2]。また、エージェントメディアエイティドコミュニケーションにおいても、エージェントの感情などの内部状態の表出である表情が効果的であることが知られている[11]。よって、擬人化エージェントをより信頼できるものにするために、感

情モデルを中心にエージェントの内部状態のモデル研究が行われてきた[2][13]。我々は、このようなエージェントの表現が起因するエージェントの内部状態を広くエージェントの**マインド**(mind)と呼ぶ。

しかしながら、マインドが十分に実装されたとしても、人間のユーザとエージェントの間で互いのマインドの解釈が困難であるという重要な問題がある。

擬人化エージェントと人間との間に効果的なコミュニケーションを実現するためには、双方が相手の表情から相手のマインドを同定できることが必要である。なぜなら、相手のマインドを考慮した行動をとることが求められるからである。例えば、人間のマインドが“忙しい”や“イライラする”であるときは、エージェントはできるだけ人間の作業を邪魔しないように振る舞うべきである。また、逆にエージェントが不安そうな表情で意見を言うとき、人間は自信のなさを感じるというように、エージェントは人間にマインドを推定させることで、言語外の情報を伝えることができる。このように、表情から相手のマインドを同定することは、人間とエージェントの自然なコミュニケーションの実現に不可欠である。

相手の表情からマインドを同定することは、良く知った人間同士のコミュニケーションでは日常的に行われている容易な行為であるが、たとえ人間同士でも初対面の相手、とりわけ文化や習慣のことなる人間同士

[†] Mutual Adaptation of Mind Mappings between a Human and a Life-like Agent

Seiji YAMADA and Tomohiro YAMAGUCHI

^{*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, Information Engineering

では難しく、ましてや人間と擬人化エージェント間では困難である。例えば、図1は、3つのMicrosoft Agentのとる7つのマインドとその表情を表している。我々日本人（少なくとも筆者ら）が比較的容易にマインドを同定できるもの(Surprised, Congratulate)もあるが、ほとんど同定できないもの(Confused, Decline, Process)もある。エージェントのマインドのモデルが十分な精度で実装されたとしても、この困難さは独立に残るものである。一方、エージェントが人間の表情からそのマインドを同定することも、これまで表情認識の研究[6]が示しているように困難な問題である。これらに対して、我々は、人間や擬人化エージェント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び、相手のマインドマッピングを相互に学習して相互適応することをマインドマッピングの相互適応と呼ぶ。

本論文では、人間と擬人化エージェントがマインドマッピングをお互いに学習して相互適応を実現する枠組みを提案する。両者は、相互読心ゲームと呼ばれるゲームを行うことにより、相手のマインドマッピングを学習していく。エージェントは、事例に基づく学習により、マインドマッピングを学習する。そして、PCとCCDカメラを用いてこの枠組みを実装し、我々の枠組みで実際にマインドマッピングの相互適応が実現されることを実験的に示す[14]。また、マインドマッピングの相互適応を高速化するヒューリスティックスを提案する。以下、まず関連研究について述べる。

Velásquez は、ミンスキーの“心の社会”に基づく感情モデルを提案している[13]。基本的な感情のエージェントからなるマルチエージェントシステムを構成し、それらのエージェント間のインタラクションの結果として感情が生成される。この研究に代表される感情モデルの研究は、感情の生成が目的であり、マインドマッピングの相互適応を目的とはしていない。

人間とコミュニケーションするアバタ、communicative エージェントの研究も行われている[3]。しかし、それらは、あくまでエージェントの設計、実装で留まっており、本研究のように人間とエージェントが相互適応

するためのインタラクション設計を行っているわけではない。

Steels は、会話と通じて2つのエージェントが語彙を学習していく区別ゲームを提案している[12]。マインドマッピングは、一種の語彙と考えられるが、2者が同一対象について会話する区別ゲームに対し、お互いが相手の表情について会話するマインドマッピングの相互適応は、学習対象である相手の内部状態が変化する可能性があるため、潜在的により困難な問題であり、我々の研究対象とSteelsらのそれは異なる。

近藤らは、エージェントメディアエイティッドコミュニケーションのためにユーザの実表情を擬人化エージェントの合成表情で置換するシステムにおいて、ユーザとのインタラクションを通じてユーザの望んでいる表情と実際の生成された表情のずれを補正する手法を提案している[7]。ユーザに直接修正された事例に基づく学習により補正が行われる。この研究は、ユーザ(人間)側のマインドから表情へのマッピングをシステム(エージェント)が適応学習していく興味深い研究であるが、本研究のような相互適応を扱っていない。

2. マインドマッピングの学習

2.1 構成要素

本節では、人間とエージェント間におけるマインドマッピングの相互適応を扱うための枠組みを定式化する。まず、提案する枠組みの構成要素を以下に示す。

- **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であり、基本マインドが代入される。 s_h, s_a はそれぞれ、人間とエージェントのマインド状態である
- **基本マインド** $E^h = \{e_1^h, \dots\}, E^a = \{e_1^a, \dots\}$: 基本マインドは、離散的で記号で記述できるとし、人間とエージェントの基本マインドの集合を E^h, E^a とする。
- **基本表情** $X^h = \{x_1^h, \dots\}, X^a = \{x_1^a, \dots\}$: X^h と X^a は、人間とエージェントが表出できる基本的な表情の集合である。一般に、一つのマインドに対し、複数の表情が表出される。
- **マインドマッピング** $M_h : x \rightarrow e = \{x_h \rightarrow e_1^h, \dots\}, M_a : x \rightarrow e = \{x_a \rightarrow e_1^a, \dots\}$: 基本表情から基本マインドへの多対1の写像。 $M_h : x \rightarrow e$ と $M_a : x \rightarrow e$ がそれぞれ人間とエージェントの真のマインドマッピングである。また、人間(エージェント)により学習されたエージェント(人間)のマインドマッピングを $M_a^h : x \rightarrow e, M_h^a : x \rightarrow e$ とする。
- **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あり、マインドマッピング

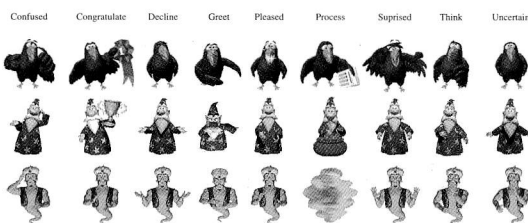


図1 擬人化エージェントのさまざまな表情

の逆写像である。

- **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ことができる。

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のように表現できる。この図において、マインドマッピングの学習とマインドマッピングの相互学習は、以下のように定義できる。

- **マインドマッピングの学習**: 人間(エージェント)がエージェント(人間)のマインドマッピング $M_n^a : x \rightarrow e$ ($M_n^h : e \rightarrow x$) を獲得すること。
- **マインドマッピングの相互学習**: 人間とエージェントがお互いにマインドマッピングの学習をすること。

2.2 エージェントの知識に対する仮定

本研究で、エージェントが知っている構成要素を設定する。エージェントは、設計される対象なので、エージェント自身の基本マインド、基本表情、マインド遷移関数を知っているとすることは自然である。

さらに、人間(ユーザ)の基本マインドも、エージェントがそれに対応して行動する必要がある基本マインドのみが重要なので、ここでは、エージェント側から必要となる人間の基本マインドを規定するアプローチをとる。つまり、エージェントとのコミュニケーションに用いる人間の基本マインドは、本来人間がもっているすべての基本マインドではなく、エージェントが利用可能な基本マインドがエージェントにより設定される。

一方、人間の学習、適応の戦略については制約を与えない。よって、人間はエージェントのマインドマッピングを好きなように学習する。

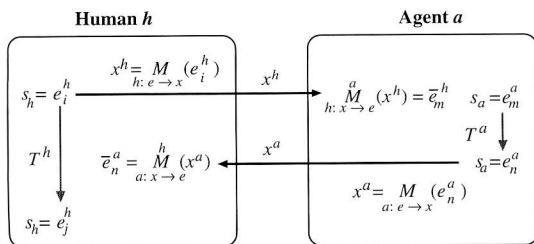


図2 人間とエージェントのインタラクションの枠組み

2.3 エージェントの学習

ここでは、エージェント側の学習について述べる。まず、人間の基本マインドが与えられるので、エージェントはそれらを学習により獲得する必要はない。また、人間の基本表情は、CCDカメラでキャプチャされる。よって、エージェントが人間の表情 x^h を観測し、それからマインドマッピングにより人間の基本マインド e^h を推定した場合に、人間によりその推定が正しいと判断されれば、 $x^h \rightarrow e^h$ を人間のマインドマッピングの事例として蓄えることができる。このように、事例を蓄えていくことにより、事例ベース学習[1]、あるいは、NN(Nearest Neighbor)法[4]により、エージェントは人間のマインドマッピングの学習を実現できる。

後述する相互読心ゲームにおいて、エージェントが人間のマインドを推定し、それを人間に提示したとき、人間は“Yes”か“No”で返答する。初期段階では、事例が少数でありエージェントの推定は間違っている場合が多いため、“No”の返答が繰り返される。一般的な事例に基づく学習では、否定事例を使わないが、ここでは否定例を積極的に利用できるようにIBL2[1]を改良した学習アルゴリズムを用いた。

否定例を扱うために、推定されたクラス(基本マインド)に対し、人間の判定が“Yes”の場合は肯定票を投票(voting)し、“No”の場合は否定票を投票する。そして、最も類似した事例最近一定数の投票において、(肯定票数)-(否定票数)が最大であるクラスを判定クラスと利用する。アルゴリズムの詳細を、図3に示す。なお、パラメータは、経験的に $n=2$, $\alpha=900$ として以降の実験を行った。

3. 相互読心ゲーム

相互読心ゲームの目的は、ユーザに認知的負荷をかけずに、事例に基づく学習のための事例を効率よく収集することである。事例は、観測された人間の表情と推定された基本マインドの対である。このゲームにおいて、人間とエージェントは交互に相手の表情からマインドを推定し、それを相手に伝え、返答をもらい、それに基づき相手のマインドマッピングを学習することにより、相互適応が実現される。相互読心ゲームは、人間とエージェントの双方が正しいマインドマッピングを獲得したときに終了するため、一種の協調ゲームになっている。また、ゲーム形式をとることにより、ユーザが楽しみながら相互適応を実現できるため、認知的負荷も少ない。

また、相互読心ゲームにより、エージェントに対する信頼(trust)と人間の動機付けが期待できる[8]。ゲームを通じて、人間とエージェントがお互い徐々に適

Agent learning procedure

- $c \in C, c = (\bar{x}^h, S_c)$: 事例.
- \bar{x}^h : 人間の表情 x^h の属性ベクトル.
- S_c : 事例 c における最近 n 個の解答ペア集合. $S_c = \{(e_1^h, good), (e_2^h, nogood), \dots\}$
- $E^h = \{e_1^h, \dots\}$: 基本マインド e^h の集合.

1. 新たな人間の表情の属性ベクトル \bar{x}^h_{new} が与えられる.
2. 属性ベクトル間の距離を計算し, \bar{x}^h_{new} に最も類似した事例 c_{sim} を決定.
3. 下式により, 基本マインド $\hat{e}^h \in E^h$ を決定.

$$\hat{e}^h \leftarrow \operatorname{argmax}_{e^h \in E^h} \sum_{s \in S_{c_{sim}}} g(e^h, s)$$

ここで, $g(e^h, s) = 1$ if $s = (e^h, good)$, $g(e^h, s) = -1$ if $s = (e^h, nogood)$ である. 初期段階で, 事例がない場合は, \hat{e}^h をランダムに決定. また, $S_{c_{sim}}$ は, c_{sim} の解答ペア集合である.

4. \hat{e}^h を人間に提示し, Yes か, No の返答を得る.
5. 返答が Yes の場合, $(\hat{e}^h, good)$ を $S_{c_{sim}}$ に追加し, $S_{c_{sim}}$ の要素数が n を越える場合は最も古い要素を削除.
6. 返答が No の場合, $(\hat{e}^h, nogood)$ を $S_{c_{sim}}$ に追加し, $S_{c_{sim}}$ の要素数が n を越える場合は最も古い要素を削除. さらに, c_{sim} と \bar{x}^h_{new} の距離がしきい値 α 以上の場合, 新たな事例 $(\bar{x}^h_{new}, \{(\hat{e}^h, nogood)\})$ を C に追加.

図3 エージェントの事例に基づく学習アルゴリズム

応することは、「人間は最初から洗練されたエージェントをモデリングしようとしなさい」という報告[10]からも、有効であると考えられる。

3.1 相互読心ゲームのルール

相互読心ゲームのルールおよび手順を以下に示す。人間とエージェント間の情報のやりとりは、GUIを通して行われる。なお、ステップ3でエージェントは、正解を返すが、ステップ6では、人間のユーザは正解を返さない。これは、ユーザにできるだけ負担をかけないためである。

1. エージェントの表情が人間に提示される。
2. 人間は、エージェントの表情からその基本マインドを推定し、それをエージェントに伝える。
3. エージェントは、“Yes” (正しい場合)と答えるか、あるいは、“No” (間違った場合)と答えて正解のマインドを人間に提示する。
4. エージェントは、人間の表情を CCD カメラで取り込む。

5. エージェントは、取り込んだ表情からマインドマッピングを使って、人間のマインドを推定して提示する。
6. 人間は、“Yes” (正しい場合)か、“No” (間違った場合)を返答する。
7. 終了条件(後述)を満たしていれば、停止。そうでなければ、(1)へ。

3.2 成功率と停止条件

一つの基本マインド e に対する成功率 $r(e)$ は、最近の成功率として下式で計算される。 S_e^g は、図3のアルゴリズムにある最近の解答ペア集合 S_c のうち、正解のペア $(e, good)$ を含む S_c の集合である。

人間 (エージェント) についてすべての基本マインド e に対する $r(e)$ の平均成功率 R により、人間とエージェントのそれぞれの学習の進捗を評価する。

$$r(e) = \frac{S_e^g \text{中の成功解答ペア総数}}{S_e \text{中の解答ペア総数}}$$

相互読心ゲームの終了条件は、人間とエージェントの双方で $R = 1$ となることである。つまり、人間とエージェントの両方で、最近一定数のマインドの推定が、すべての基本マインドにおいて正しく行われることである。

4. 実装

ノート PC (SONY VAIO-SR9G/K) と USB 接続のカラー CCD カメラ (Creative Media: WebCam Plus, 解像度 720×680 , 8bit カラー) を使って、提案した枠組みを実装した。VineLinux 2.5 上で Gtk+, Video4Linux API を使い、C 言語でプログラミングした。実験環境を図4に示す。

4.1 事例の表現と取り込み

エージェントの学習のときには、エージェントは 500 ms 毎に人間の表情を取り込み、連続 4 回の取り込み画像間の差異がしきい値以下になった場合に、その画像を事例とする。つまり、人間の表情が 2 秒間のストップモーションになったときに、その表情を取り込む。このしきい値は、後述するユークリッド距離で 250 と経験的に決めた。

このストップモーションによるキューにより、人間は表情を設定しやすくなる。

取り込まれた画像は、解像度 40×30 の 8bit グレイスケールに変換されて、1200次元の 256値をとる多次元ベクトルの事例として蓄えられる。この変換は、事例間の類似度計算を高速にする意味がある。また、予備実

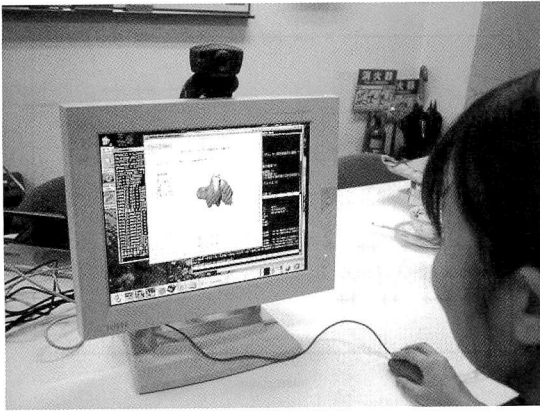


図4 実験環境

験において、この程度の解像度のグレイスケールでも妥当な結果が得られている。また、類似度には、ベクトル間のユークリッド距離を使った。

なお、本研究では、取り込んだ画像のピクセルを事例の次元と考へ、通常の表情認識で用いられる特徴抽出や領域分割を行っていない。本研究では、表情認識の実現が目的ではなく、エージェントの表情認識能力に応じて人間が適応することに目的があるので、敢えてそのような画像処理をやっていない。その結果、後述する実験において、人間がエージェントの認識能力に適応する現象が観察されている。

4.2 ユーザインタフェース

図5は、人間がエージェントの表情からそのマインドを推定しているときのGUIのスナップショットである。人間が“Start”ボタンをクリックすると、エージェントの表情が提示される。そして、人間はエージェントのマインドを推定し、そのマインドを“Primitive emotions”ボタンの中から選んでクリックする。エージェントがユーザの推定したマインドを判定しているのが、図6である。ウィンドウの下にあるプログレスバーが、人間とエージェントの平均成功率 R を表しており、両方のバーが右端まで達した場合 ($R=1$) にゲームが終了する。人間は、プログレスバーにより、お互いの学習の進捗を知ることができる。

図7は、エージェントが人間のマインドを推定しているGUIのスナップショットである。人間が“Start agent's recognition”ボタンをクリックすることで、エージェントによる人間の表情の取り込みが始まる。安定した表情を取り込んだ後、その4つの画像がウィンドウに表示される。また、これまでに事例として蓄えられた各画像と今取り込まれた画像との距離が表示されている。

エージェントは、2.3のアルゴリズムにより、人間の

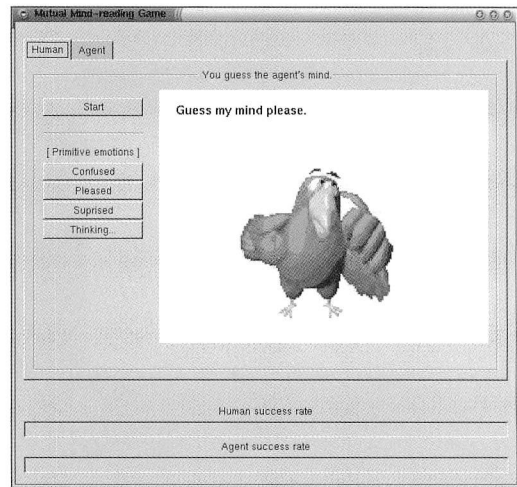


図5 人間がエージェントのマインドを推定

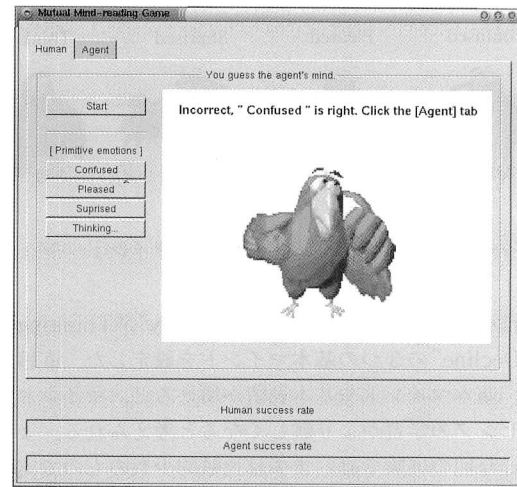


図6 人間の推定したマインドをエージェントが判定

マインドを推定し、それを図7のように人間に提示する。人間は、それに対し、“Yes”か“No”のボタンをクリックして判定する。

5. 評価実験

我々の枠組みで、人間とエージェントの相互適応が実現されることを確認するため、実験を行った。すべての実験において、情報系大学院学生5名とスタッフ3名（うち、女性は1名のみ）が被験者である。

図8にあるような4つの基本マインドと基本表情をエージェントに設定した。この4つの基本マインドでは、“Confused”と“Think”の区別が難しく、人間がエージェントのマインドマッピングを学習するのはそ

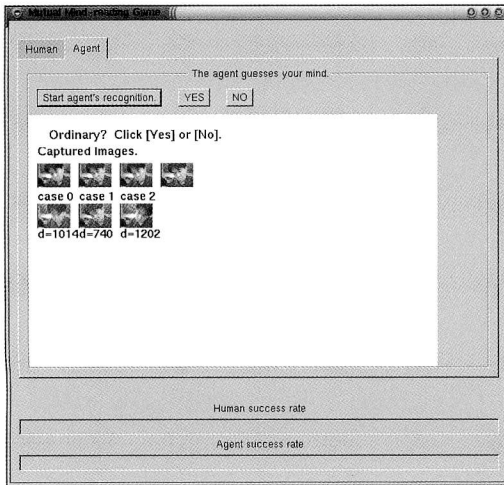


図7 エージェントが人間のマインドを推定

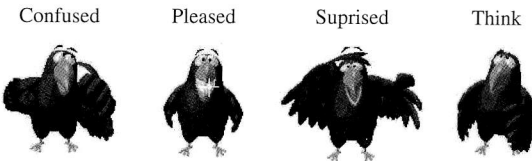


図8 エージェントの基本マインド

れほど容易ではない。人間には“Ordinary”, “Thinking”, “Decline”の3つの基本マインドを設定した。直感的に、基本マインドや基本表情が増えると、マインドマッピングの学習がより難しくなると考えられる。よって、今回の実験では、まずは比較的少数のもので実験をし、相互適応が収束することを調べることにした。

実験に先立ち、被験者に以下のような簡単な説明を与えた。しかし、エージェントの学習アルゴリズム、平均成功確率の計算式、表示された人間の画像の意味などの詳細は教えていない。

- 相互読心ゲームのルール。
- GUI についての最小限の説明。
- 人間の表情の取り方へのアドバイス：あまり微妙な表情の違いは認識できない旨を説明した。
- 人間の取り得る3つの基本マインド。

また、2.1で述べたエージェントのマインド遷移関数 $T^a(c)$ は、マインドがランダムに遷移するように設定した。このマインド遷移関数の改良については、後述する。一方、人間側のマインド遷移については、インタラクションを与えずに、自由にしてもらった。

以上の条件で、8名の被験者それぞれに対して、エ

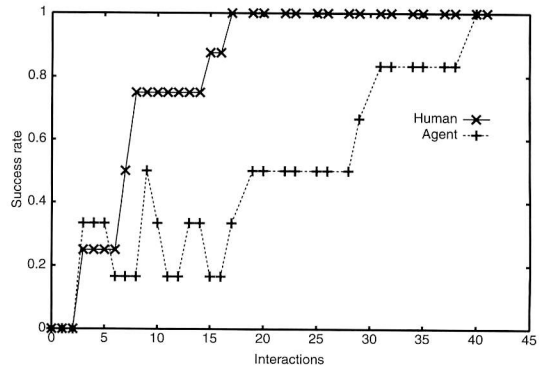


図9 マインドマッピングの相互適応

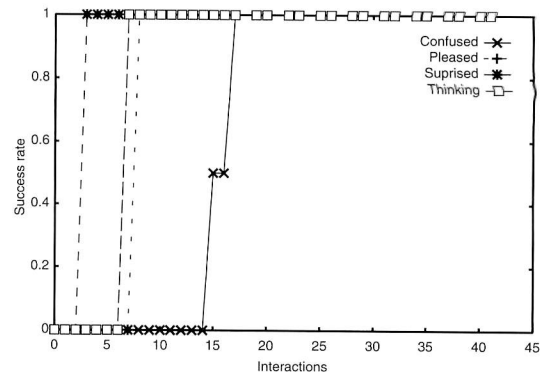


図10 人間によるエージェントの基本マインド毎の学習

ージェントと相互読心ゲームをやってもらい、人間とエージェントの(平均)成功率の推移、ゲーム終了までのインタラクション回数、経過時間を調べた。なお、人間とエージェントとの一回づつのマインド推定を、1インタラクションとする。

5.1 相互適応の観察

図9は、人間とエージェントの平均成功率の典型的な推移を示す。双方の平均成功率は、インタラクションを重ねるにつれ徐々に増加していき、最後にはゲームが終了する。人間もエージェントも時折マインドの推定に失敗するので、成功率は単調増加にはならない。他のすべての被験者においてほぼ同様の推移が見られた。これにより、被験者全員で、人間とエージェントの相互適応が実現されていることが確認された。なお、一回のゲームに要する時間は、5~10分であり、後で行ったアンケートではほとんどの被験者がゲームを楽しんだと答えている。

ある被験者におけるエージェントの基本マインド毎の成功率の推移を、図10に示す。このグラフから、“Confused”と“Think”などの判別が難しい表情においても、最終的にはうまく学習されていることがわかる。この傾向は、他の被験者でも見られた。よって、本実験のような少数の基本マインド（基本表情）においては、人間は高いマインドマッピングの学習能力を示すことがわかる。

図11は、各被験者(s-1, ..., s-8)について、人間の学習とエージェントの学習が収束するまでのインタラクション数を示している。これを見ると、まずこの実験においては、人間の学習効率がエージェントのそれよりもよいことがわかる。また、人間の学習はほぼ20インタラクション前後で収束しているが、エージェントの学習の方は大きくばらついている。これは、被験者によっては、エージェントの学習の能力では、判別できない些細な変化の表情を長く、あるいは何度も提示していることによる。

獲得された事例の例として、図12に3名の被験者s-1, s-2, s-4の事例の画像を示す。被験者s-2が最小の3つの事例なのに対し、s-1, s-4は、5つ以上の事例を獲得している。すべての被験者においても、獲得される事例数にはばらつきがある。また、図12からわかるように、多くの被験者が、エージェントの表情認識能力が高くないという認識から、頭を傾けたり、手を顔にかざしたりすることで、表情の差異を強調していることがわかる。

6. エージェントの戦略的マインド遷移関数

さらに、マインド遷移関数を改良した実験を行った。これまで、エージェントのマインド遷移をランダムに行っていたが、それでは、人間側がマインドマッピングを学習済みの基本マインドにエージェントが遷移することがあり、無駄が多い。よって、最も成功率の悪い基本マインドに決定論的に遷移するマインド遷移関数を適用し、ランダムなマインド遷移関数との比較実験を行った。

前節での実験において、エージェントのマインド遷移関数のみを、学習率最悪の基本マインドに遷移するものに変更して、同じ被験者で実験を行った。エージェントのマインド遷移関数は、主に人間の学習効率に影響を与えられるので、人間の平均成功率が1になるまでのインタラクション回数を調べた。なお、実際の実験は、各被験者で、最初に成功率による遷移関数で実験を行い、次にランダム遷移による実験を行っている。よって、ゲームを経験することによる学習効果は、成功率による遷移関数にとって有利には働いていない。

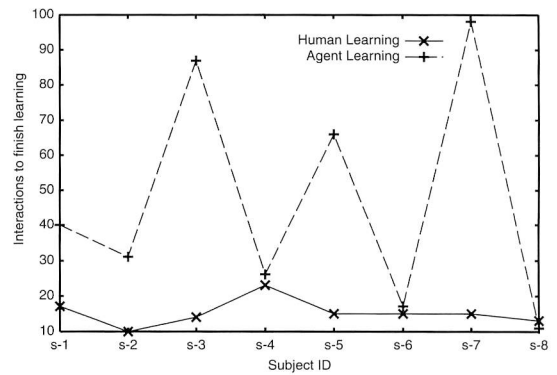


図11 学習にかかったインタラクション数

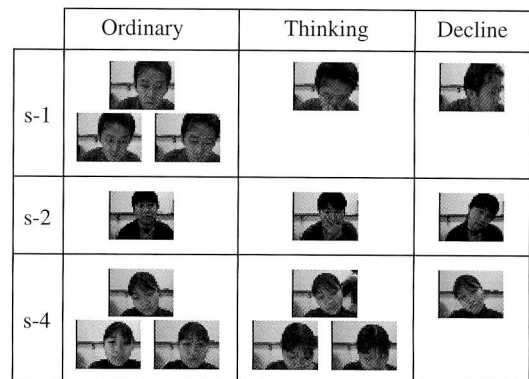


図12 いくつかの事例の画像

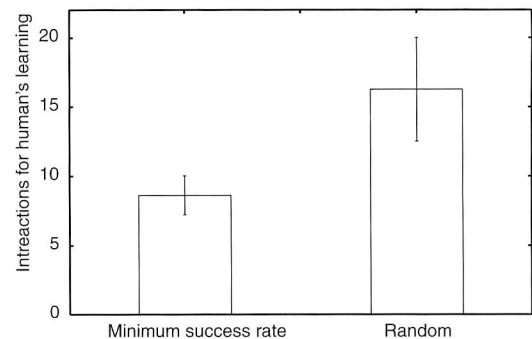


図13 戦略的なマインド遷移関数

ていない。

実験結果を図13に示す。これは、各被験者の平均成功率が1になるまでに必要だったインタラクション数の平均を示しており、 t -検定 ($\alpha=0.05$) で有意差があった。このように、成功率によるマインド遷移関数は、ランダムな遷移関数よりも、人間の学習効率を向上さ

せることがわかった。

一方、現在の枠組みでは、エージェントのマインド遷移、表情マッピングは、事前に設計されたものに固定されている。今後は、人間とのインタラクションを通じて、個々の人間の嗜好を考慮しながら、適応的にマインド遷移、表情マッピングを変化させることが重要であろう。

7. 課題

7.1 相互読心ゲームと日常の作業とのギャップ

実験結果でも明らかなように、実際のマインドと必ずしも一致しない表情をとっても、相互読心ゲームを終了できるため、相互読心ゲームにおいて人間は誇張した表情をとる傾向にある。マインドマッピングは、日常的な作業におけるユーザの表情からマインドを推定するために利用されるので、このような相互読心ゲームにおける表情と日常の表情のギャップは問題である。よって、相互読心ゲームのルールに、人間はマインドと一致した自然な表情をとることなどを追加することを検討している。

7.2 相互読心ゲームのメリット

相互読心ゲームでは、交互に相手のマインドを当て合っているが、交互にやる必要性はあるのか、つまりエージェント(人間)の学習を収束するまでやり、その後人間(エージェント)の学習をやるというように、順番に独立してやるのと何が違うのかという疑問がある。

これに対して、まず交互に学習することが楽しいこと、次にゲームをすることで相手のマインド遷移を推定することが容易になり、相互適応が加速化されること[16]などが、相互読心ゲームを行うメリットとして挙げられる。

7.3 相互適応の促進

6節の戦略的マインド遷移により、人間の適応は促進された。では、エージェントの適応は促進されていないのだろうか。実は、実験を通して、人間側が誇張した表情(7.1節)や表情マッピングの一貫性と保つなどの、エージェントの学習を促進するための行動が観察されている。このように、相互適応を促進するように、ユーザインタフェース、学習アルゴリズムを設計すること、つまり、相互適応を促進するインタラクション設計が重要である。そのためには、本研究では固定されているエージェントの表情マッピング、基本表情などを人間の嗜好に適応させる必要があると考えられる。また、今回の実験は、相互適応を確認するに留

まっており、基本マインド/表情、表情マッピングなどのパラメータを変えた組織的な実験も必要である。

8. まとめ

人間のユーザと擬人化エージェントがお互いの表情の解釈を学習し、相互に適応する必要を示し、そのようなマインドマッピングの相互適応を実現する枠組みを提案した。人間とエージェントが、互いに相手の表情からマインドを推定し合う相互読心ゲームを構成し、相互適応が実現されることを実験により示した。さらに、人間側の適応を促進するためのマインド遷移関数を開発し、その効果を実験的に示した。

現在、擬人化エージェントをペットロボットに置き換え、人間とロボットとの相互適応についても研究を進めている[15]。人間が大型ペットロボットを犬としてモデル化することを利用したヒューリスティクスを開発、検証している。

謝辞

本研究遂行にあたり、文部科学省科学研究補助金基盤(C)(課題番号:14580410)の補助を受けました。記して感謝いたします。

参考文献

- [1] D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms. *Machine Learning*, Vol. 6, pp. 37-66, 1991.
- [2] J. Bates. The role of emotion in believable agents. *Communications of the ACM*, Vol. 37, No. 7, pp. 122-125, 1994.
- [3] J. Cassell. Embodied conversational agents: Representation and intelligence in user interface. *AI Magazine*, Vol. 22, No. 4, pp. 67-83, 2001.
- [4] B. V. Dasarathy. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. IEEE Computer Society Press, 1991.
- [5] 石塚満. マルチモーダル擬人化エージェントシステム. システム/制御/情報, Vol.44, No.3, pp.128-135, 2000.
- [6] H. Kobayashi and F. Hara. Recognition of six basic facial expressions and their strength by neural network. In *IEEE International Workshop on Robot and Human Communication*, pp. 381-386, 1992.
- [7] 近藤崇, 角所考, 美濃導彦. 対話的に獲得される事例に基づく行為者指向の顔メディア変換. システム制御情報学会論文誌, Vol. 14, pp. 308-315, 2000.
- [8] P. Maes. Agents that reduce work and information overload. *Communications of the ACM*, Vol. 37, No. 7, pp. 30-40, July 1994.
- [9] MS agent Web page. <http://msdn.microsoft.com/msagent/>.
- [10] B. A. Myers, A. Cypher, D. Maulsby, D. C. Smith, and B. Shneiderman. Demonstrational interfaces.

- In *Proceedings of 1991 Conference on Human Factors and Computing Systems*, pp. 393-396, 1991.
- [11] K. Rivera, N. J. Cooke, and J. A. Bauhs. The effects of emotional icons on remote communication. In *CHI'96 companion*, pp. 99-100, 1996.
- [12] L. Steels. Emergent adaptive lexicon. In *Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pp. 562-567, 1996.
- [13] J. D. Velásquez. Modeling emotions and other motivations in synthetic agents. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, pp. 10-15, 1997.
- [14] S. Yamada and T. Yamaguchi. 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.
- [15] S. Yamada and T. Yamaguchi. Training AIBO like a dog. In *The 13th IEEE International Workshop on Robot-Human Interaction*, pp. 431-436, 2004.
- [16] T. Yamaguchi, Y. Ochi, and S. Yamada. Mutual adaptation in communication between a human and a life-like agent by sharing context dependent facial expressions. In *4th Asia-Pacific Conference on Simulated Evolution And Learning*, 2002.
(2004年10月31日 受付)
(2004年12月25日 採録)

[連絡先]

〒101-8430 東京都千代田区一ツ橋2-1-2

国立情報学研究所知能システム研究系

山田 誠二

Tel&Fax : 03(4212)2562

E-mail : seiji@nii.ac.jp

著者紹介



やま だ せい じ
山田 誠二 [非会員]

1984年 大阪大学基礎工学部卒業, 1989年 同大学院博士課程修了, 同年大阪大学基礎工学部助手, 1991年 同大学産業科学研究科講師, 1996年 東京工業大学大学院総合理工学研究科助教授, 2002年 国立情報学研究所教授, 現在にいたる, 工学博士, 人工知能, 特に, 知的 Web, ヒューマンエージェントインタラクションに興味をもつ。



やま ぐち とも ひろ
山口 智浩 [非会員]

1985年 大阪大学基礎工学部制御工学科卒業, 1987年 同大学院基礎工学研究科修士課程修了, 同年 三菱電機(株), 1988年 松下電器産業(株)を経て1991年 大阪大学基礎工学部助手, 1998年 奈良工業高等専門学校情報工学科助教授, 現在に至る, 強化学習法を基にマルチエージェント学習, ヒューマンエージェントインタラクション, ユーザ間インタラクション支援の研究に従事, 2000年 第14回人工知能学会全国大会優秀論文賞, 人工知能学会, 情報処理学会, 計測自動制御学会, ロボット学会など各会員, 工学博士。

Mutual Adaptation of Mind Mappings between a Human and a Life-like Agent

by

Seiji YAMADA and Tomohiro YAMAGUCHI

Abstract :

This paper describes a human-agent interaction in which a user and a life-like agent mutually acquire the other's mind mapping through a mutual mind reading game. In these several years, a lot of studies have been done on a life-like agent such as a Microsoft agent, an interface agent. Through development of various life-like agents, a mind like emotion, processing load has been recognized to play an important role in making them believable to a user. For establishing effective and natural communication between an agent and a user, they need to read the other's mind from expressions and we call the mapping from expressions to mind states *mind mapping*. If an agent and a user don't obtain these mind mappings, they can not utilize behaviors which significantly depend on the other's mind. We formalize such mutual mind reading and propose a framework in which a user and a life-like agent mutually acquire each other's mind mapping. In our framework, a user plays a mutual mind reading game with an agent and they gradually learn to read the other's mind through the game. Eventually we fully implement our framework and make experiments to investigate its effectiveness.

Keywords : Mutual adaptation, Life-like agent, Human-Agent Interaction

Contact Address : **Seiji YAMADA**

National Institute of Informatics

2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-8430, JAPAN

TEL : 03(4212)2562

FAX : 03(4212)2562

e-mail : seiji@nii.ac.jp