

人間-エージェント間での読心ゲームにおける

言い当て行動系列に依存したユーザモデルの推定

Estimating a User Model with the Action Sequences in the Mind Reading Game
between a Human and a Life-like Agent山口 智浩^{*1}
Tomohiro Yamaguchi上利 宗久^{*1}
Munehisa Agari山田 誠二^{*2}
Seiji Yamada^{*1} 国立奈良工業高等専門学校
Nara National College of Technology^{*2} 国立情報学研究所
National Institute of Informatics

This paper proposes an adaptive interaction between a human and a life-like agent. We present the method to estimate the partner's context with the n th order MDP model. In this framework, an agent plays a mind reading game with a user in which the agent estimates the user's current mind state by the previous game context. Then the agent gradually learns to read the user's mind transition rules by estimating the transition probabilities of the n th order MDP model. To examine the regularity of the partner's context, we made the simple experiments to estimate a user's mind transition rules. As the experimental result, the number of correct guess of four subjects is much improved in the latter half of the game

1. はじめに

本論文では、人間とエージェントとのインタラクションにおいて、人間のモデル同定を行うエージェントについて述べる。一般にインタラクションにおける相手モデル同定は一方的ではなく、特に人間や学習エージェントを対象とする場合、必ず相互作用を伴う性質がある。我々はこういう問題を人間とエージェントとのインタラクション **HAI** (Human Agent Interaction) における**相互適応 HAMA** (Human Agent Mutual Adaptation) と呼ぶ。HAMA とは、エージェント側が人間に適應してパーソナライズする過程で、人間が暗黙、意識的両面でエージェントに適應し、結果として両者が互いに適應する状況を指す[山田 03]。

HAMA では、本来両者の協調的かつ適切な適應が理想であるが、従来手法においては、エージェント側の不適切な適應が人間(以下、ユーザと呼ぶ)側の適應に悪影響を及ぼし、ユーザに負荷を与える問題点があった。そこで我々は、人間-エージェント間での読心ゲームを対象タスクとして、表情とマインドとの対応を表すマインドマッピングを両者が同時並列的に学習する枠組を提案[Yamada 02]し、両者の学習が干渉しない条件について議論し、学習実験を行なった[山口 02] [Yamaguchi 02]。その結果、ユーザの柔軟かつ強力な学習能力に対して、エージェントの学習能力が限定されかつ弱いことがわかり、これが理想的な相互適応に対してボトルネックのひとつとなっていることが示唆された。このエージェントの学習能力の限界のひとつとして、マインドマッピングの学習しか行なわず、エージェントのマインド遷移規則が組み込みであり、繰り返しゲーム中でユーザのマインド遷移の規則性を学習しない点があった。

そこで本論文では、エージェント側の適應能力を改善することを目的として、人間-エージェント間での繰り返し読心ゲームにおけるマインド遷移規則のモデルのオンライン同定とモデル利用によるゲームの言い当て精度の改善について述べる。本研究では、エージェントがユーザのマインドの言い当てを繰り返す過程で、マインド遷移系列を利用したユーザのマインド遷移の規則性のオンライン同定を行う。既存研究に対する本研究の位

置付けは、本手法によるマインド遷移規則のモデル同定をマインドマッピングの学習 [山口 02]と組み合わせてインタラクション全体の適應能力および精度を向上させることである。

2. 読心ゲームにおけるユーザモデルの同定

2.1 Human Agent Interaction のモデル化

本節では、Human Agent Interaction におけるユーザ、エージェント共通のモデルについて説明する。まず両者はそれぞれ、内部状態とそれを遷移させる規則を持つと仮定する。内部状態を**マインド(mind)**、**状況**(2.3 節で後述)に応じてマインドを遷移させる規則のことを**マインド遷移規則**と呼ぶ。マインドは外部から直接観測不能であるとし、マインドに依存して観測可能な状態のことを**表情(expression)**と呼ぶ。表情に対応するマインドへの写像のことを**マインドマッピング**[Yamada 02]と呼ぶ。本研究で扱うモデルでは、同定対象の相手モデルに対して以下を仮定する。

- マインド、表情の集合は、それぞれ有限かつ既知。
- マインドマッピングは固定かつ個性的。
- マインド遷移規則は、個性的。

ここで**個性的**とは、ユーザ、エージェントごとに異なり、共通とは限らない性質で、自己のは既知、相手のは未知であると仮定する。なお、本論文ではマインド遷移規則の同定を目的とし、表情およびマインドマッピングは扱わない。

2.2 繰り返し読心ゲーム

まず、読心ゲームについて説明する。**読心ゲーム(Mind Reading Game)**とは、状況を共有する2者 A, B が状況の変化に応じて遷移した相手のマインドを推定して言い当て、その言い当てに対する結果を教示してもらうゲームである。図1にゲーム1回の手続きを示す。

繰り返し読心ゲームとは、図1の手続きを一定回数繰り返すゲームのことである。本実験における状況とは、このゲームにおける言い当て結果(の系列)に依存すると仮定する。読心ゲームでは、マインド遷移規則の同定に対してゲームを行う両者が協調的、競争的双方の場合が考えられるが、本論文では協調的な場合を対象とする。

- step1 マインド遷移: B は, 直前の状況に応じて, あるマインドに遷移する.
- step2 言い当て: A は, B のマインドを推定し, マインド名を相手に提示する.
- step3 結果の判定と教示: B は, 提示されたマインドが自己のマインドと一致すれば Yes を, 一致しなければ No を返し, 合わせて自己のマインド名を相手に教示する.

図1 読心ゲームの手続き

2.3 読心ゲームにおけるマインド遷移規則

読心ゲームにおけるマインド遷移規則について説明する. まず, 本論文で用いる 3 種類のマインド名と省略記号との対応を図2に示す. マインド遷移規則とは, 知覚した状況を入力とし, 遷移するマインドを出力とする規則である. 繰り返し読心ゲームにおける**状況**とは, ゲームにおける言い当て結果(の系列)である. **言い当て結果**とは, 言い当て直前のマインドと相手から言い当てられたマインドとの組み合わせとし, t 回前の言い当て結果を r_t で表す. 読心ゲームにおける単純なマインド遷移規則の例を図3に示す. 図3(a)は, 直前のマインド(Normal)と一致するマインド(ここでは Normal)の言い当て r_1 に対し, Pleased マインドに遷移する規則を表す. 図3(b)は, 直前のマインド(Normal)と異なる(ここでは Pleased)言い当て r_1 に対し, Confused マインドに遷移する規則を表す. ここで**単純**とは, 状況が系列ではなく, 直前($t=1$)の言い当て結果のみに依存することを表す. 次に単純でないマインド遷移規則の例を図3(c)に示す. この場合, 状況が言い当て結果の系列 r_2, r_1 で表されている. この例では, 2 回前のマインド(Normal)が正しく言い当てられて Pleased マインドに遷移した後1回前の Pleased マインドも正しく言い当てられて, Pleased マインドに遷移する規則を表す.

C: Confused, N: Normal, P: Pleased

図2 用いる 3 種類のマインド名と記号との対応

- $r_1(N,N) \rightarrow P$... (a)
 - $r_1(N,P) \rightarrow C$... (b)
 - $r_2(N,N), r_1(P,P) \rightarrow P$... (c)
- r_t (直前のマインド, 相手の言い当て) \rightarrow 遷移マインド

図3 読心ゲームにおけるマインド遷移規則の例

2.4 n 重 MDP モデルによるマインド遷移規則の表現

n 重 MDP モデルによるマインド遷移規則のモデル化について説明する. まず, MDP (マルコフ決定過程) モデルとは, マルコフ性が仮定される確率的な状態遷移を表すモデルであり, 状態をノード, 行為をアークとする確率的状態遷移グラフで表される. **マルコフ性**(厳密には一様単純離散マルコフ過程)とは, 離散時間的に状態遷移が起こり, 状態遷移確率が一定(時間的に一様)かつ直前の状態遷移のみ(単純)で決定可能な性質である. つまり現在の状態と行為の遷移確率とが分かれば, 次の状態を推測できる. **n 重マルコフ性**とは, 状態遷移確率が n 回前までの状態遷移系列のみに依存して決まる性質であり, $n=1$ の場合を**単純**と呼ぶ. n 重 MDP モデルは, n 回前までの状態遷移を状態記述に取り込めば, 単純 MDP モデルに変換される.

では, 単純 MDP モデルによる単純なマインド遷移規則のモデル化について説明する. 単純なマインド遷移規則は, 自己のマインドを状態, 相手の言い当てを行為とする単純 MDP モデルでモデル化できる. 図4に, 図3(a)の $r_1(N,N) \rightarrow P$ のモデル化の例を示す.

次に単純でないマインド遷移規則のモデル化について説明する. 単純でないマインド遷移規則では, 入力となる状況が言い当て結果系列となる. 系列を無視して単純 MDP モデルでモデル化すると, 直前の言い当て結果が同一でも, それ以前の言い当て結果系列が異なる遷移規則が存在する場合, 遷移先が確率的に枝分かれして観測される. これは直前以前の言い当て結果系列を考慮しないために生じる部分観測問題である. これを回避するため, n 回前までの(自己の)マインド遷移系列を**文脈(context)**として状態に取り込み, 状態を文脈ごとに分割することで部分観測状態を解消する.(但し, 本論文では相手の言い当てマインドは文脈に含めない. 理由は, モデルおよび同定の精度を犠牲にしてコストを小さくするためである.)これにより, 部分観測における確率的な状態遷移が決定的に近い状態遷移でモデル化されると期待できる. 図5に例を示す. 図5(a)にマインド遷移規則を示す. *は don't care 記号を表す. 図5(b)は, 直前の言い当て結果 $r_1(P,P)$ に対して, 前々回の言い当て結果 r_2 を文脈として取りこむことで状態が PP, NP, CP の3状態に分割され, N と P とに確率的に枝分かれしていた遷移がそれぞれ異なる決定的な状態遷移でモデル化されることを示している.

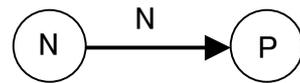


図4 単純なマインド遷移規則のモデル化

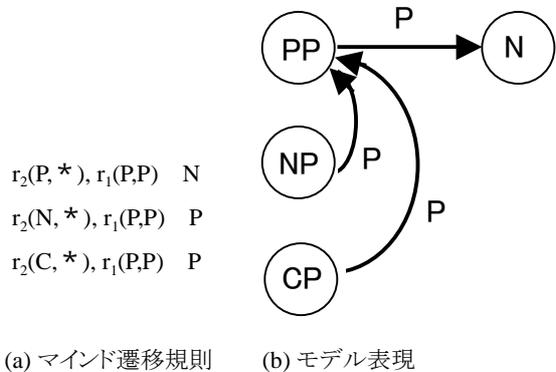


図5 単純でないマインド遷移規則のモデル化

2.5 n 重 MDP モデルによるマインド遷移規則の同定

n 重 MDP モデルによるインタラクションする相手のマインド遷移規則の同定について述べる. まず, 相手のマインド遷移規則同定の目的について述べる. 読心ゲームにおける相手モデル同定の目的は, 遷移する相手マインドの言い当てを正確に行うことである. 読心ゲームにおける相手マインドの言い当ては決定的に行なわれるため, 同定する相手モデルは, 確率的よりは, 正確かつ決定的なモデルの方が言い当て精度が向上する. したがって, モデル同定の目標は, 正確かつ決定的な相手のマインド遷移規則をモデル同定することである.

では, 単純なマインド遷移規則の同定について述べる. 相手モデル同定の目的は, 全ての状態遷移における遷移確率を最

尤推定することである。統計学で、最尤推定確率は単に出現頻度で求めた確率と等しいことが知られているので、ある状態 i 、行為 a ルールにおける遷移先状態 j への遷移確率 $P(i, a, j)$ は、ルールの生起頻度 $N(i, a)$ に対する遷移先状態の生起頻度 $N(i, a, j)$ を用いて、以下の式(1)で逐次的に推定できる。

$$P(i, a, j) = N(i, a, j) / N(i, a) \quad \dots(1)$$

次に単純でないマインド規則の同定について述べる。問題は、必要最小限の長さの文脈を決定することである。状態分割に用いる文脈長が短いと、モデルは小さくて済むが状態遷移、すなわち言い当てが確率的となる。これに対し、状態分割に用いる文脈長を長くすると、状態遷移確率が1または0に近づき、言い当てが決定的になる反面、モデルが大きくなって、空間計算量が增大するだけでなくモデル同定に要するサンプリング数、すなわち学習コストが増大する。

したがって、理想的にはオンライン時に適切な文脈の長さを推定することが望ましい。ヒューリスティックな手法として、ゲームの文脈(言い当て結果系列)を全て記憶しておき、各状態で用いる文脈長 n をパラメータ(初期値は $n=1$, 単純)として持ち、状態遷移確率が決定的でないルールを持つ状態に対して、状態遷移確率が決定的となるまで n を増加させる方法が考えられる。この手法を**可変長文脈**と呼ぶ。これに対し、扱う文脈の長さを固定する方法を**固定長文脈**と呼ぶ。

3. 実験

3.1 実験タスク

2章で説明した繰り返し読心ゲームをタスクとして、人間-エージェント間において、ユーザのマインドをエージェントが言い当てながらユーザのマインド遷移規則のモデル同定を行う2つの実験について述べる。実験1では、エージェントの言い当て戦略をランダムに固定してユーザのモデル同定を行い、実験前半と後半とでユーザのマインド遷移モデルがどの程度一貫しているかを検証する。実験2では、エージェントの言い当て戦略を実験前半はランダムに固定してユーザのモデル同定を行い、実験後半では、前半で同定したマインド遷移モデルを用いて推定したマインドへと言い当て戦略を変更して、読心ゲームの正答率を調べ、オンラインモデル同定の評価を行う。

3.2 実験方法

まず2つの実験に共通する実験条件について述べる。被験者は、研究室の学生4名、教官1名の計5名である。ユーザが取りうるマインドは 2.3 節、図2の3種類、ゲームの繰り返し回数は前半15回、後半15回の計30回(マインド当りの期待試行回数は5回*2)である。同定に用いるモデルは、2.4, 2.5 節で説明した文脈を用いる MDP モデルである。

次に実験で用いたエージェントの**言い当て戦略**について述べる。尚、両戦略共に、毎回言い当て結果を用いて 2.5 節の方法でモデルを逐次更新する。

- **ランダム戦略**: 状況に関係なくランダムなマインドを提示する。モデル同定を行う探索戦略としては最も単純な行動選択戦略である。
- **モデル推定戦略**: オンラインで同定したモデルを参照して、直面した状況に対し最も遷移確率の大きいマインドを提示する。(但し未探索の状態の場合には、ランダム戦略を用いる。)

測定項目は、実験前半、後半それぞれの、ユーザのマインド遷移規則の同定モデルおよび読心ゲームの正答数である。次に各実験ごとの設定条件を説明する。

(1) 実験1: ユーザのマインド遷移規則の一貫性の評価実験

- 実験の目的: 実験前半と後半との言い当て結果を用いて同定した2つのユーザモデルを比較し、ユーザが用いたマインド遷移規則の一貫性を評価する。
- エージェントの言い当て戦略: ランダム戦略
- モデルの文脈長: 長さ2の固定長文脈
- 被験者へのインストラクション: "あなたの感じた気持ちに一番近いマインドを1つ選んで入力してください。"

(2) 実験2: オンラインモデル同定を用いた言い当て実験

- 実験の目的: 実験前半でオンライン同定したユーザのマインド遷移モデルを用いて、後半で言い当ての推定を行い、実験後半での読心ゲームの正答率がどの程度改善されるかを評価する。併せて、言い当て戦略変更に伴う被験者の印象の変化についてアンケート調査を行う。
- エージェントの言い当て戦略: 前半はランダム戦略、後半はモデル推定戦略
- モデルの文脈長: 初期長1の可変長文脈
- 被験者へのインストラクション: "あなたの感じた気持ちに一番近いマインドを1つ選んで入力してください。エージェントが正確に言い当てられるように協調してください。"

3.3 実験結果

(1) 実験1: ユーザのマインド遷移規則の一貫性の評価実験

図6に5名の被験者のうち、3名の被験者の実験前半と後半とのマインド遷移規則の同定結果を示す。但し、状態遷移は各ルールの枝分かれ確率の最も大きい遷移のみを表示している。まず、3名それぞれのマインド遷移規則がある程度異なることから、マインド遷移規則は**個性的**であるといえる。また、ゲームで用いられ、探索されたマインド遷移規則はモデル全体に対して**部分**

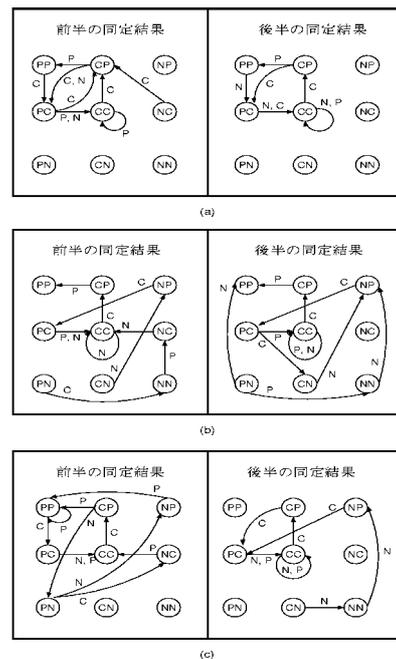


図6 実験1: ユーザのマインド遷移規則の同定結果

的である。これは、マインド数3、文脈長2の状態遷移ルール数が $3^2 \times 3 = 27$ であるのに対し、実験でのサンプリング数が前半、後半それぞれ15回と半分程度であるのが主たる原因である。また同一の被験者において用いられたマインド遷移規則は、実験の前半と後半とで一部異なっており、時間的にも部分的であることがわかる。

次に実験前半、後半でのユーザのマインド遷移規則の一貫性について述べる。図6の結果から、3名の被験者に関しては、同一のルールで(最大遷移確率の)状態遷移先が異なるルールがひとつもないことから、実験1においてユーザの**マインド遷移規則は一貫**していると考えられる。なお、大半のマインド遷移規則は**決定的**であり、少数の確率的ルールは、ユーザの規則性のゆらぎを表すと考えられる。

(2) 実験2:オンラインモデル同定を用いた言い当て実験

図7に、5名の被験者それぞれの実験前半と後半での言い当て成功回数の比較を示す。5名中4名の被験者で、実験後半の言い当て正答数が増加している。マインド数3におけるランダム戦略による正答率の期待値は $1/3$ であり、これは実験前半15回の言い当てにおける正答数の平均値 5.6 とほぼ一致するのに対し、実験後半の平均正答数は 10.8 回と大きく上回っている。したがって、今回実験した協調タスクの場合には、オンラインモデル同定を用いてタスク後半の言い当て精度が改善可能だといえる。

次に実験後の被験者アンケート結果について述べる。意図的に規則性を持たせて入力したと答えた被験者が1名(C)、実験前半でこちらの意図を汲んでくれず、正答率の低さに閉口した、と答えた被験者が2名(B,D)いた。うち1名(D)は、実験後半で相手の反応が読めてきたので協調できた、と回答した。これらは図7の実験結果の解釈と一致するといえる。

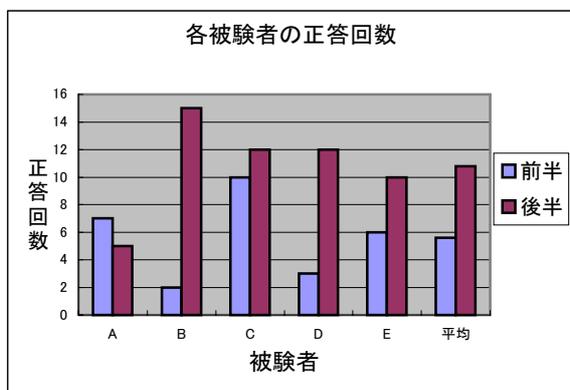


図7 実験2:実験前半と後半での言い当て成功回数の比較

3.4 考察

(1) オンラインモデル同定における探査とタスク達成効率

3.3 節、実験2の結果について考察する。図7において興味深いのは、前半2番目に正答数の多かった被験者 A の後半の正答数が最も少なく、一方、前半最も正答数の少なかった被験者2名(B,D)の後半の正答数が上位を占めたことである。前者の場合、被験者 A のモデル同定結果を分析したところ、実験の前半と後半とでモデルの一貫性はあったが、前半のランダム探査によるモデル同定が部分的であり、後半で未探査のマインド遷移規則に多く直面したためにランダム言い当てが多くなり、後

半の正答数がランダム戦略と近くなっていることがわかった。これに対し後者の場合、前半でのランダム探査によってモデル中の多くのルールの同定が行えたため、後半の言い当ての正答率が大幅に向上したといえる。これらに対し、前半後半を通して正答数が多かった被験者 C の場合には、ゲームを通して用いた(探査した)ルール数が少なかったためと考えられる。以上より、繰り返し数30回という短期的な条件では、モデル同定において同定の探査範囲の広さとタスク達成効率(本実験では正答率)とが両立しにくいといえる。

(2) 人間ユーザのモデル同定における適切な探査戦略

実験1のモデル同定の結果から、事前の予想に反してユーザのマインド遷移規則は決定的に近いことがわかった。さらに実験2の被験者アンケートから、ランダム戦略による探査はゲームの正答率が悪いと、ユーザにストレスを与えていることがわかった。統計的には生起確率が不変な場合、サンプリング数が大きい程、生起確率の推定精度が向上するが、人間相手の場合には、多くのサンプリングによって生じるストレスが、マインド遷移規則の一貫性に悪影響を与えることが予想される。したがって、人間ユーザのモデル同定における適切な探査戦略は、強化学習における **k-確実法** のように各ルールの探査回数 k を一定に保ち、かつ $k=1$ に近い小さい値で十分であると予想される。

4. おわりに

本論文では、人間-エージェント間での繰り返し読心ゲームにおける、マインド遷移系列を利用したマインド遷移規則のオンラインモデル同定とその利用について述べた。実験1の結果から、ゲームで用いられたユーザのマインド遷移規則には一貫性があり、個性的、決定的な性質が強く、モデル全体の網羅的探査に対してユーザの用いる規則は部分的であることがわかり、実験2の結果から、ゲーム中にオンライン同定したモデルを用いて、ユーザに外乱を与えることなく言い当て精度が改善されることがわかった。今後の課題は、モデル同定の適切な探査戦略を導入することと、本手法によるマインド遷移規則のモデル同定を読心ゲームタスクにおけるマインドマッピングの学習と組み合わせ、学習精度を向上させることである。

参考文献

- [山口 02] 山口, 越智, 山田: 文脈依存表情の共有によるヒューマンエージェント間コミュニケーションでの相互適応, 第 16 回人工知能学会全国大会, 2B3-07, 2002
- [Yamaguchi 02] T.Yamaguchi, Y.Ochi, S.Yamada: Mutual adaptation in communication between a human and a life-like agent by sharing context dependent facial expressions, 4th Asia-Pacific Conference on Simulated Evolution And Learning (Seal2002), #2192, 2002
- [Yamada 02] S. Yamada, T. Yamaguchi: Mutual Learning of Mind Reading between a Human and a Life-like Agent, *The 5th Pacific Rim International Workshop on Multi-Agents (PRIMA-2002)*, K.Kuwabara and J.Lee (Eds.), Lecture notes in Artificial Intelligence 2417, Springer, pp.138-150, 2002
- [山田 03] 山田, 山口: 人間とエージェントの相互適応, 第 16 回人工知能学会全国大会, 2D1-02, 2003