

人間－エージェント間での言い当てゲームにおける相手モデルの同定

Identification of partner's model in the Mind Reading Game between a Human and a Life-like Agent

大西 浩司*¹
Koji Ohnishi

山口 智浩*¹
Tomohiro Yamaguchi

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 presents an adaptive interaction between a human and a life-like agent. We present the method to estimate the partner's context with the nth 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 nth order MDP model. In our previous research, we showed an ability of the nth order MDP model to identify user's model. However, there is a major problem that the random estimation by the agent makes user's mind transitions unnatural. To identify users' natural mind transitions, natural interaction is essential. Therefore, we propose a new action selection method using built-in averaged model during on-line model identification. We show the simple experiment compared our method with random action selection method, then we discuss the characteristics of our proposed method.

1. はじめに

学習において、エージェントが相互作用する環境や、相手にモデルを仮定し、相手モデルを同定することは重要である。強化学習におけるモデル同定では、できるだけ広い範囲を効率よく同定するためのモデル探索のための行動選択手法としてランダム探索法や ϵ -確実法などが提案されてきた。しかしながら、静的な環境や固定アルゴリズムによるエージェントと異なり、学習適応能力を持つエージェントや人間ユーザの場合、従来手法ではモデル同定のための幅広い探索が相手の規則性を変動させようという新たな問題が生じる。特に後者の場合、人間が普段用いない遷移に対して探索を繰り返すことで、人間が相手に対する不自然さを感じ取り、それによって人間の行動モデルが変動するという問題が示された[山口 2003]。

そこで本論文では、相手に不自然さを感じさせない範囲内での幅広いモデル同定を行うための手法として、あらかじめ同定した複数の被験者から作成した**組み込み平均モデル**を用いた探索によるモデル同定手法を提案する。前回の実験で用いたランダム行動によるモデル同定手法との比較実験を行うことで、それぞれの手法の特徴を分析し、提案手法の問題点と改善のアイデアについて議論する。

2. 言い当てゲームにおける相手モデル同定

2.1 言い当てゲーム

本研究のエージェントは**言い当てゲーム**を繰り返し行うことでユーザに適応する。このゲームでは、ユーザとエージェントは相手に対するモデルとして、外部から直接観測できない**内部状態**である**マインド**と、**状況**に応じてマインドを遷移させる規則である**マインド遷移規則**を持つと仮定する[山口 2003]。言い当てゲームは、状況に応じて遷移するユーザのマインドをエージェントが言い当てるゲームである。本研究では言い当てゲームを通じた

マインド遷移規則の同定を目的に被験者実験を行う。

2.2 マインド遷移規則

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する規則である。言い当て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文脈である。エージェントが言い当てゲームの文脈から相手のマインド遷移規則を学習することをモデル同定と呼ぶ。なお、マインドには「Confused (困った)」、「Normal (普通)」、「Pleased (嬉しい)」の3種類を用いる。

2.3 言い当てゲームの手順

次に言い当てゲームの手順を図1に示す。まず状況に応じてユーザがマインドを遷移させる(Step1)。次にエージェントがマインドを推定し、相手マインドの言い当てを行う(Step2)。最後にエージェントからの言い当てに対して、ユーザが正しいマインドをエージェントに教示する(Step3)。ユーザからの教示と言い当てが一致したなら言い当て正答である。

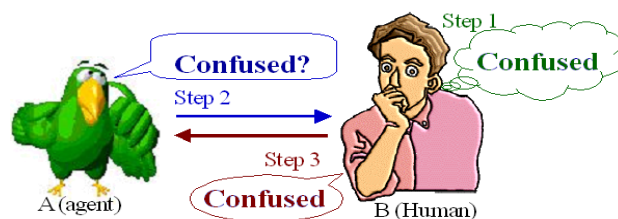


図1 言い当てゲームの手順

2.4 本研究におけるモデル同定の目標

本研究における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数多くモデル同定すること」である。まず、正確に同定する理由について述べる。言い当てゲームでは同定したモデルを基にユーザのマインドを推定する。エージェントが様々な状況に対処するため、マインド遷移規則を多く、正確にモデル同定する必要がある。

連絡先: 大西浩司, 奈良工業高等専門学校, 〒639-1080
奈良県 大和郡山市 矢田町 22, 0743-55-6140,
Mail:yamaguch@info.nara-k.ac.jp

次に決定的に同定する理由について述べる。言い当てゲームにおける相手マインドの言い当ては決定的に行われるため、確率的なモデルで言い当てるよりも、状況を細分化した決定的なモデルを用いて決定的に言い当てる方が言い当て精度が向上するからである。

2.5 モデル同定で扱う言い当てゲームの文脈

言い当てゲームで扱う文脈について述べる。状態遷移を対象とするモデル同定において、モデルで扱う文脈が直前の状態遷移に限定する場合は**単純なモデル**、それ以外の長さ2以上の文脈を扱う場合は**複雑なモデル**と呼ぶ。複雑であるほど長い文脈を扱うため、詳細にモデル同定できる反面、同定に要するサンプリング数や計算量が増大する。前回の実験より、単純なモデルよりも長さ2の複雑なモデルの方がモデル同定の目標を達成していることが明らかになっている。そこで今回の実験では長さ2のモデルをモデル同定に用いる。

2.6 マインド遷移図

言い当てゲームによって得られたユーザのマインド遷移とエージェントからの言い当ての系列を表現した確率状態遷移グラフを**マインド遷移図**と呼ぶ。マインド遷移図における3種類のマインド名と省略記号との対応を図2に示す。

C: Confused, N: Normal, P: Pleased

図2 用いる3種類のマインド名と記号の対応

今回用いる複雑なモデルでは、マインド遷移図における状態は長さ2のマインド遷移系列で表す。図3に長さ2のマインド遷移規則の例を示す。矢印前後の○は言い当て前後の状態を、矢印上の記号は行動を表す言い当てマインドである。図3を用いてマインド遷移図で用いる記号を説明する。

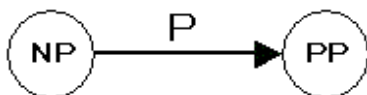


図3 長さ2のマインド遷移規則の例

マインド遷移図では、図3のように各状態を矢印で結ぶことでマインドの遷移を表す。状態「NP」は、NからPへと遷移を行い、現在、すなわち言い当て直前のマインドがPであることを示す。遷移元の状態は NP、遷移先の状態はPPとなっており、エージェントからのPという言い当てによってマインドが P に遷移したことを表す。

図4に複雑なモデル全体のマインド遷移図を示す。マインド数3、文脈長2なので、状態数は $3^2=9$ である。図4の①のPNは、ユーザのマインドがPからCへ遷移した状態を示している。

②の矢印はマインド遷移規則を表現しており、ユーザの状態がCNであるとき、エージェントが現在のマインドをNと言い当てた結果、ユーザのマインドがNからPに変化し、状態がNPへと遷移したことを表現している。状態CNにおけるユーザの言い当て直前のマインドはNである。エージェントからの言い当てはNであり、現在のユーザのマインドと一致するので、この場合の**言い当て正答**となる。エージェントからの言い当てが正答した回数

を**正答回数**と呼ぶ。言い当て正答時のマインド遷移規則を**正答規則**といい、正答規則の個数を**正答規則数**と呼ぶ。

③のマインド遷移規則は、現状態NPから次状態PPへ遷移する際、確率的な枝分かれが無い決定的なマインド遷移を行っている。このような、確率的な枝分かれが無いマインド遷移規則を**決定的規則**といい、その個数を**決定的規則数**と呼ぶ。

図4のマインド遷移図の状態CNにおけるマインド遷移規則数は2個、NPにおけるマインド遷移規則は2個存在するので、図4のマインド遷移図には**探索規則数**が4個存在するといえる。

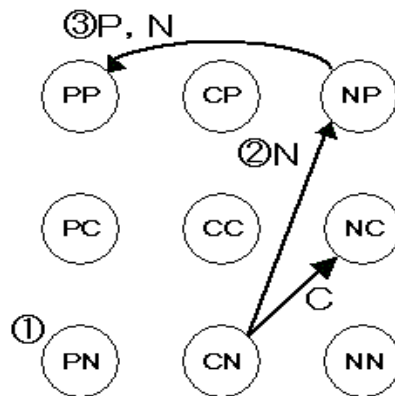


図4 複雑なモデルを用いた場合のマインド遷移図

2.7 組み込み平均モデルを用いた探索

組み込み平均モデルを用いた探索について述べる。従来は相手のモデルを学習するために、オンライン同定の前半では相手のマインドをランダムに言い当てていた。しかしながら、ランダム言い当ては広い範囲のマインド遷移規則を学習できる反面、文脈の一貫性のない不自然な言い当てがユーザへのストレスとなり、ユーザの規則性が変化することが前回の実験で示唆された[山口 2003]。

そこで本論文では、オンライン同定の前半での組み込み平均モデルを用いた探索法を提案する。前回の実験で得られた11名分のマインド遷移系列を用いて学習したモデルを**組み込み平均モデル**と呼ぶ。このモデルを用いた決定的言い当てを行うことで、被験者11人から得られた平均的な文脈に沿った探索が行われるため、自然な言い当てを実現できると考えられる。

3章の実験では、オンライン同定の前半での探索の比較条件として従来のランダム言い当てと、組み込み平均モデルを用いた決定的言い当てのどちらの方が自然な言い当てであるかを明らかにする。

3. 実験

本章では、人間—エージェント間での言い当てゲームにおいて、ランダム言い当てと組み込みモデルを用いた決定的言い当てのどちらが、ユーザにとって自然な言い当てが可能であるか明らかにするため行った実験について述べる。

3.1 実験方法

(1) 実験条件

言い当てゲームにおける比較条件:

言い当てゲームでは、エージェントからの言い当てを30回行う。総回数30回のうち前半15回の言い当てを比較条件としてオ

ンライン同定を行い、後半15回では、前半に同定したユーザのモデルを用いてエージェントがユーザのマインドの推定を行いながら、オンラインモデル同定を行う。比較条件は、ランダム言い当てと組み込み平均モデルを用いた決定的言い当てである。

被験者:

本校の情報工学科本科5年生と専攻科の男子学生らの計22名の被験者を、実験前半15回に組み込みモデルを用いた言い当てを行うグループ(以下「Gb1」)11名と、ランダム言い当てを行うグループ(以下「Gr2」)11名とに分けて実験を行う。

実験環境:

言い当てゲームでは、被験者と言い当てゲームのプログラムとがペアとなり実験を行う。実験は被験者に外乱が加わらないよう、1人ずつ個室で行う。

(2) 評価基準

自然な言い当てを実現するためには、不自然な言い当てによる言い当てタスクの達成効率の低下を防ぎつつ、2.4節で述べたモデル同定の目標を達成することが必要である。そこで今回の実験では、言い当てタスクの達成効率を下げずにモデル同定の目標を達成することを中心に実験結果を評価する。評価に用いる行動的指標[小松 2003]として、**平均正答回数**、**平均探査規則数**、**平均決定的規則数**、**平均正答規則数**の4つを用いる。

まず**平均正答回数**は、言い当てタスクの達成効率の評価を表す。平均正答回数は被験者グループでの言い当て正答回数の平均で、言い当てタスクの達成効率を示している。この指標が大きいほど言い当てタスクの達成効率が良いことがわかる。

残りの3つの指標である**平均探査規則数**、**平均決定的規則数**、**平均正答規則数**は、モデル同定の目標の達成についての評価を表す。**平均探査規則数**は被験者グループでの同定したマインド遷移規則数の平均で、ユーザのモデルでの同定した範囲を示している。この指標が大きいほど広い範囲でモデル同定したことになる。**平均決定的規則数**は被験者グループでの決定的規則数の平均で、指標が大きいほど決定的なモデル同定を示す。**平均正答規則数**は被験者グループの正答規則数の平均で、この指標が大きいほど正確なモデル同定を示す。

モデル同定の目標は前述のとおり「正確かつ決定的な相手のマインド遷移規則を数多くモデル同定すること」である。モデル同定の目標において、正確なモデル同定には平均正答規則数を、決定的なモデル同定には平均決定的規則数を、相手マインド遷移規則のモデル同定の広さは平均探査規則数を用いて評価を行う。これらの指標が大きいほどモデル同定の目標を達成していることを示す。

3.2 実験結果

Gb1の実験結果を表1に、Gr2の実験結果を表2に示す。両者の比較結果として、表3にGb1とGr2との差分を示す。結果をまとめると、組み込み平均モデルGb1では、前半の言い当てはランダム探査Gr2より良いが、前半でのモデル探査能力がランダム探査に劣るため、後半の正答回数の改善量が小さいと言える。詳細は各評価項目で述べる。

平均正答回数:

表3より、平均正答回数を比較する。ゲーム全体では、Gb1とGr2はほぼ同じ回数の言い当てに正答していた。また、前半か

ら後半にかけての回数の変化はGb1が+1.3、Gr2が+3.3とともに増加している。このことから、ゲーム前半で同定したユーザのモデルが、後半の言い当てに適切に反映されたことがわかる。

しかしながら、ゲーム前半ではGr2よりGb1の方が多く、ゲーム後半では逆にGb1よりGr2の方が多い。つまり、組み込み平均モデルGb1では、前半の言い当てはランダム探査Gr2より良いが、改善量が小さいため、後半の正答回数ではランダム探査に逆転されていることがわかる。

表1 Gb1の実験結果

	前半	後半	全体
平均正答回数(回)	6.2	7.5	13.6
平均探査規則数(個)	6.7	9.5	13.3
平均決定的規則数(個)	2.9	3.9	2.9
平均正答規則数(個)	1.8	3.4	4.1

表2 Gr2の実験結果

	前半	後半	全体
平均正答回数(回)	4.9	8.2	13.1
平均探査規則数(個)	9.1	7.7	13.8
平均決定的規則数(個)	4.1	4.0	4.8
平均正答規則数(個)	3.1	2.6	4.9

表3 Gb1とGr2の差分(Gb1-Gr2)

	前半	後半	全体
平均正答回数(回)	1.3	-0.7	0.5
平均探査規則数(個)	-2.4	1.8	-0.5
平均決定的規則数(個)	-1.2	-0.1	-1.9
平均正答規則数(個)	-1.3	0.8	-0.8

平均探査規則数:

表3より、平均探査規則数は、ゲーム全体としてはGb1とGr2はほぼ同じ個数の規則を探索しているが細かく見ると、ゲーム前半ではGb1よりGr2の方が多く、ゲーム後半では逆にGr2よりGb1の方が多い。これはゲーム前半ではGr2の方が広い範囲でユーザがマインド遷移を行い、ゲーム後半ではGb1の方が広い範囲でユーザがマインド遷移を行ったことを示している。

また、Gb1は前半から後半にかけて探査規則数が増加、逆にGr2は前半から後半にかけて規則数が減少していることがわかる。このことから、Gb1は前半から後半にかけてマインド遷移の範囲が広くなり、逆にGr2は前半から後半にかけてマインド遷移の幅が狭くなったことがわかる。つまり組み込み平均モデルGb1では前半でのモデル探査能力がランダム探査Gr2に劣るため、後半に平均モデル以外の遷移に多く遭遇すると言える。

平均決定的規則数:

表3より、ゲーム全体としては、Gr2の方がGb1より平均で1.9個多いことから、同じゲーム回数に対してGr2の方がGb1より多くの決定的遷移を同定したことがわかる。また、表1、2より、Gb1は前半から後半にかけて平均決定的規則数が増加するが、Gr2はほぼ同じである。次に表3より、平均決定的規則数は、ゲーム前半ではGb1よりGr2の方が多く、ゲーム後半ではほぼ同じ個数になっている。これは、ゲーム前半ではGr2の方が決定

的な遷移を行うことが多く、ゲーム後半ではGb1とGr2でほぼ同じ個数の決定的遷移規則を用いることを示している。これらの差異の原因は言い当て方法の違いである。Gb1では前半の、組み込み平均モデルに基づく決定的言い当てが被験者に合わない場合に正答数と決定的規則数が少なくなるのに対し、Gr2は前半全員に対しランダム言い当てを行うため約 1/3 の割合で言い当てが正答する。一方、後半は両者とも同定モデルに基づいて同じ方法で言い当てを行うため傾向に差が出なくなる。

平均正答規則数:

表1と表2より、平均正答規則数は、ゲーム前半ではGr2の方がGb1より多く、ゲーム後半ではGb1の方がGr2よりやや多くなっている。これは、ゲーム前半においてはGb1よりGr2の方が効率よく正答規則を同定し、ゲーム後半においてはGr2よりGb1の方が効率よく正答規則を同定していることがわかる。ゲーム全体としては、Gr2の方がGb1より 0.8 個多いことから、同じゲーム回数に対してGr2の方がGb1よりやや多くの正答規則数を同定したことがわかる。表1と表2より、Gr2は前半から後半にかけて平均正答規則数が減少しているのに対し、逆にGb1は平均正答規則数が増加している。これらの原因として、Gb1は前半に局所的な遷移を行う傾向があり、一方Gr2は後半に少数の正答規則間での局所的な遷移を行う傾向が強いことが示唆される。

3.3 考察

(1) モデル同定の目標達成の効率

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数多くモデル同定すること」である。正確なモデル同定は平均正答回数と平均正答規則数、決定的なモデル同定は平均決定的規則数、同定したマインド遷移規則の多さは平均探索規則数をそれぞれ示している。

まず正確なモデル同定について議論する。Gb1とGr2のゲーム全体の平均正答回数はほぼ同じ値であったが、ゲーム全体の平均正答規則数に関してはGr2の方がやや多い。このことからGr2の方がGb1より正確なモデル同定を行うことがわかる。しかしながら、ゲーム後半に関しては、平均探索規則数と平均正答規則数は、Gb1の方がGr2よりも多い。このことは、ゲーム後半では、Gb1の方がGr2より正確なモデル同定が可能なことが示唆される。

次に決定的なモデル同定について議論する。実験結果より、ゲーム全体の決定的規則数はGb1よりGr2の方が多いことがわかる。このことから、Gr2の方がGb1より決定的なモデル同定を行うことがわかる。また、ゲーム前半については、Gr2の方がGb1より多かったが、ゲーム後半にはGb1とGr2はほぼ同じ個数、決定的な遷移を同定できたことがわかる。

最後に同定したマインド遷移規則の個数について議論する。実験結果より、ゲーム全体の平均探索規則数はGb1とGr2はほぼ同じ値だった。このことから、Gb1とGr2に対して、ほぼ同じ範囲で探索を行ったことがわかる。しかしながら、Gb1の後半における平均探索規則数はGr2よりも多く、ゲーム後半においてはGb1の方がより多くのマインド遷移規則の探索を行っていることがわかる。

以上より、今回の実験ではゲーム全体を通じてはGb1よりGr2の方がよりモデル同定の目標を達成しているといえる。

(2) 自然な言い当てについての比較

言い当てゲームにおける自然な言い当てについて議論する。自然な言い当てを行うためには、相手の行動モデルを変動させ

ない範囲内での、より広い範囲の同定と正しい言い当てが必要である。また、エージェントからユーザに対する適応のしやすさ (adaptive) だけでなく、ユーザがシステムの挙動を適宜制御できること (adaptable) が必要である。

前回の実験においてランダム言い当ては文脈に依存しない不自然な言い当てが多く、被験者の行動モデルを変動させることが示唆されていた。しかしながら、ランダム言い当ては正答回数、探索規則数、決定的規則数、正答規則数について、被験者間で大きな差が無い。つまり、全ての被験者に対して、ある程度適切なモデル同定が可能であることがわかる。このことから、ランダム言い当てによるモデル同定は adaptive であるが adaptable で無いことがわかる。

これに対して、Gb1の組み込みモデルを用いた決定的な言い当てによるモデル同定は被験者間の正答回数、探索規則数、決定的規則数、正答規則数の差は大きい。総被験者の半分がGr2の平均正答回数より多くなっている。このことから、被験者によっては、組み込みモデルを用いた決定的な言い当てはランダム言い当てより適切なモデル同定を行なったことがわかる。しかしながら、単一の組み込みモデルによる決定的な言い当てでは全ての被験者に対応できないことも示唆される。つまり、組み込みモデルを用いた決定的な言い当てによるモデル同定は adaptable であるが adaptive で無いことがわかる。

これらのことにより、自然な言い当てを可能にするモデル同定を実現するためには、ランダムな言い当てを用いたモデル同定の adaptive と、組み込みモデルを用いた決定的な言い当てによる adaptable を両立する必要があると考えられる。これを実現する方法として、組み込みモデルによる確率的な言い当てによって平均モデルに沿った確率的行動探索を行い、adaptive と adaptable の両立を図ることを検討している。これにより、組み込みモデルを用いた決定的な言い当てによるモデル同定と、ランダムな言い当てを用いたモデル同定の中間的な結果が得られることが期待される。

4. 結論

本論文ではユーザとエージェントとの言い当てゲームにおける相手モデル同定について、自然な言い当てを実現するための組み込み平均モデルを用いた探索による相手モデル同定手法を提案し、ランダム探索法との比較実験を行い、提案手法の特徴、問題点と改善のアイデアについて述べた。今後の予定は、組み込み平均モデルを用いた確率的行動探索による比較条件を加えて本論文で実験した2手法との比較を行い、人間にとって自然な言い当てとは何かを明らかにすることである。

参考文献

- [山口 2003] 山口智浩, 大西浩司, 山田誠二: 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, pp.345-350, JAWS2003, 2003.
- [小松 2003] 小松孝徳, 宇都宮淳, 植田一博, 岡夏樹: 人間-エージェント間における「自然なコミュニケーション状態」の客観的指標, pp.351-357, JAWS2003, 2003.
- [角 1999] 角康之: 情報可視化システムにおける適応的インタラクション, 人工知能学会誌, pp.33-40, オーム社, 1999.