

# 2本腕バンディット問題に対する 「緩い対称性モデル」の有効性： 因果推論における対称性バイアスと相互排他性バイアス

篠原 修二\*

中野 昌宏†

2007年3月24–25日 進化経済学会第11回大会（於・京都大学）

## 要 旨

本報告は、因果関係に関する「人間らしい」推論の1つのモデルを提案する。

従来より、 $p$  という事象と  $q$  という事象とのあいだにどのような関係があればそれが「因果関係」と認知されるのか、について議論されてきた。われわれは、その認知の中で働いていると見られる「対称性バイアス」および「相互対称性バイアス」という2つの非論理的な認知的バイアスが、結果として合理的な因果関係の認知・推論・期待形成をもたらしているのではないかという仮説を立てた。

これを検証するため、この2つのバイアスが柔軟に働く「緩い対称性モデル」を定義し、コンピュータ内に構築したエージェントに実装して、このエージェントが「2本腕バンディット問題」に対して示すパフォーマンスを調べた。その結果、このエージェントの振る舞いは人間による被験者実験のデータときわめてよく合致するだけでなく、機械的な硬直性を超えた柔軟かつ人間らしい挙動を示すことがわかった。

このように、「緩い対称性モデル」は人間の因果推論を非常によく表現するモデルであると言える。と同時に、「対称性バイアス」と「相互排他性バイアス」は人間の推論の基礎にある本質的な契機だと考えられる。

## はじめに

古来から、哲学、心理学、精神分析学、経済学、脳科学、サイバネティクス、認知科学、人工知能研究、ロボット研究などなどの分野では、人間の行動・判断・推理のモデルがさまざまに構想されてきた。特にそこで問題とされたのは、「人間の振る舞いは動物とどう違うのか」「機械とどう違うのか」という点であろう。

たとえば、行動アルゴリズムの大部分が本能としてプログラムされている動物と異なり、人間は「本能の壊れた存在」とよく言われる [11]。他方で、近代以降「人間はよくできた機械だ」という意見 古くはデカルトやラ・メトリー、新しくはノーバート・ウィーナーなどが説得的に展開されてきたが、人間の行動原理は通常の意味での機械的なアルゴリズムだけでは描きえないということも、たとえば経済学において [21]、あるいは AI の哲学において [16] 指摘されてきた。

機械的なアルゴリズムはどうしても硬直的な行動しか導かないだけでなく、一挙手一投足について膨大な、物理的に不可能なまでの計算量を要求する。そこでこうした流れの中から、アルゴリズムに対して「ヒューリスティックス」を対置し、より簡便な方法で判断・意思決定するモデルを考案することで、人間行動を説明し

\* 豊橋技術科学大学大学院工学研究科 (PD) shinohara@vox.tutkie.tut.ac.jp

† 大分大学経済学部 nakano@cc.oita-u.ac.jp

ようとする認知心理学的着想も発展してきた。が、それによっても決定的な仮説がもたらされるまでには至っていない。

人間は、つねに論理的に推論しているわけではなく、しばしば「誤った」推論をする。が、その誤り方もランダムというわけではなく、一定の傾向性がある。そして、結果として生存戦略上「合理的」ないし「適応的」な推論と行動を帰結しているのである [8, 17, 19, 20]。本報告は、この誤り方の形式を、正しい推論にかかる一種の認知的バイアスとして取り出し、これを実現するモデルをコンピューター内のエージェントに実装することで、真に「人間的」と言える推論形式の本質に迫ろうと試みるものである。

以下、まず第1節でモデルの構築に際する基本的な考え方について議論し、次に第2節でモデルの定義とそれ自体のもつ性質について解説する。最後に第3節でそのモデルのパフォーマンスを検証する実験の枠組みと結果、そしてそれについての考察を述べたい。

## 1 モデル構築の考え方

### 1.1 論理推論と因果推論

認知心理学などの分野でしばしば指摘されることだが、純粋に論理的な推論（ここでは「論理推論」と呼ぶ）と日常的な推論とのあいだにはずれがある [24]。

たとえば、子どもの親は子どもに、「もし部屋を片づけなかったら、遊びに連れて行かないよ ( $\bar{p} \rightarrow \bar{q}$ )<sup>\*1</sup>」などによく言う [6]。が、子どもはこの言葉を「もし部屋を片づけたら、遊びに連れて行ってあげる ( $p \rightarrow q$ )」とか、「遊びに連れて行ってもらうためには、部屋を片づけなければならない ( $q \rightarrow p$ )」とか、要するに「もし部屋を片づけたら、そのときだけ遊びに連れて行ってあげる ( $p \leftrightarrow q$ )」という意味に受け取る。この場合、親の実際の意図も子どもの解釈どおりであり、親は子どもにより強い動機づけにより片づけをさせるために、方便として「 $\bar{p} \rightarrow \bar{q}$ 」という言い方を採用したのである。ともかく、一般に日常的にこのようなものの言い方が通用するということから見れば、「 $p \rightarrow q$ 」や「 $q \rightarrow p$ 」や「 $\bar{p} \rightarrow \bar{q}$ 」や「 $p \leftrightarrow q$ 」という形式の論理は、現にしばしば混同されていることになる。

他方、論理的には、これらはそれぞれ別の言明であり、混同は誤りである。冒頭の親の言葉を額面どおり受け取るかぎりには、「もし部屋を片づけたとしても、遊びに連れて行かない」という可能性も残っている。一般には逆に、こちらの理屈のほうこそ通らないだろう。

子どもは、「部屋を片づけるかどうか」という自分の行動 ( $p, \bar{p}$ ) を、「遊びに連れ出してもらえるかどうか」という親の行動 ( $q, \bar{q}$ ) に影響を与える原因として位置づけるのであり、端的に両者の行動をそれぞれ独立した事象とは捉えない。その意味で、子どもはこの推論を論理的にはなく、因果的に行っていると言えるのかもしれない。こうした日常生活上実用的な因果に関わる推論（以下「因果推論」と呼ぶ）においては、論理的意味で正確な推論は逆に有効でない場合があると言える。

### 1.2 2つの推論バイアス

対称性バイアス 多くの研究が示すところでは、人間には「 $p$ ならば $q$ である」という情報を「 $p$ である場合、かつそのときにのみ $q$ である」と、論理的に誤って受け取ってしまう傾向がある [22, 25, 6, 7, 26, 27, 28]。これは命題「 $p \rightarrow q$ 」からその逆命題「 $q \rightarrow p$ 」をも同時に想起してしまうことによる。この命題から逆命題

\*1 論理的否定を表すためには一般に「 $\neg p$ 」という表記を使うことが多いが、以下では簡略化のため「 $\bar{p}$ 」という表記で統一する。

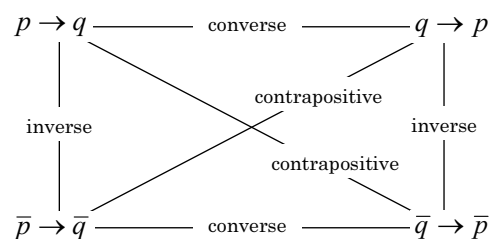


図1 逆・裏・対偶

を想起してしまうという人間の傾向は、従来からよく指摘されていることであるが、われわれはこれを「対称性バイアス symmetry bias」と名づけることにする。

これはまさに因果関係に関する推論についてかかることの多いバイアスだと言える。たとえば「雨が降ったら、地面が濡れる」ということから、「地面が濡れているのであれば、それは雨が降ったからだろう」と推測することにあたる。このかぎりでは、この推論形式は論理的には妥当ではないが、単に推論の誤りないし非論理的推論とすべきでもなく、日常生活の範囲では十分に実用的な推論形式でもあることに注目したい。

相互排他性バイアス 一方、幼児の語彙学習の研究分野では、命題「 $p \rightarrow q$ 」からその裏命題「 $\bar{p} \rightarrow \bar{q}$ 」を想起・認知する傾向が指摘されてきた。この傾向は従来より「相互排他性バイアス mutual exclusivity [bias]」と呼ばれている [12, 9, 4, 1]。

この「誤り」は幼児が諸対象の名前を学習する際には逆に有効に働く。ある対象（便宜上  $o$  と呼ぶ）の名前が  $l$  であるとすでに学んだ幼児は、別の対象（ $o'$  と呼ぶ）を見せられたときには、それがまた同じ  $l$  という名前だとはもはや思わず、それには何か別の名前があるだろうと推測する。そのことによって、正しく  $l'$  という新たな名前が覚えられるのである。しかしながら、実は  $o \rightarrow l$  であることと  $o' \rightarrow l$  であることは論理的に両立的であり、 $o'$  もが  $l$  という名前でない保証はない。実際、形状の多少異なる対象が同じ名前で呼ばれることも多々あり、幼児はその母語の中でどこまでが同じ名前で呼ばれるのかを結局は試行錯誤で学んでゆくことになる。が、当初はこのバイアスにしたがって、つまり「別の対象には別の名称がある」という仮定のもとで、幼児の学習は進んでゆくのである。

従来これらのバイアスは、それぞれ全く異なる研究分野で独立して扱われるのみで、総合的に取り扱われてはこなかった。しかしながら、これら2つの推論バイアスは、論理的には相互に密接に関連しているはずである。少なくとも命題論理の観点から見れば、「 $q \rightarrow p$ 」と「 $\bar{p} \rightarrow \bar{q}$ 」は相互に対偶関係にある [17]。したがって、 $(p \rightarrow q) \implies (q \rightarrow p)$  が言えるのであれば、自動的に  $(p \rightarrow q) \implies (\bar{p} \rightarrow \bar{q})$  も言えることになり、その逆も真である。

また、以上の説明からも見て取れるように、この両バイアスは異なる意味ではあれ、いずれも「 $p \rightarrow q$ 」なる論理を「 $p \leftrightarrow q$ 」という一対一対応（全単射）の方に引っ張る・歪めるバイアスと見ることができる。このように、いずれの観点からも、両バイアスは相互に非常に近いものであることは疑いない。

とはいえそれでも、これらのバイアスは相互に意味を異にしており、独立に定義できる。それゆえ両バイアスは論理的に同値とまでは言えず、したがって同じものというわけではない。両バイアスは、相互に非常に似ているとともに、異なったバイアスである。とすれば、これらの両方のバイアスがかかる状況では、場合に

よってはバイアスどうしが相互に干渉し、強めあったり打ち消しあったりと、複雑な効果をもたらす可能性がある。

### 1.3 因果推論と仮説形成的推論

これが推論，判断，意思決定，学習に好ましい効果を与えうはずだという見通しについては，先にあげた例，すなわち「雨が降ったら，地面が濡れる」うんぬんという推論形式の理解が傍証となるだろう。これはシャーロック・ホームズが『緋色の研究』において「逆向きの推理」と呼んだものであり [2]，誤った三段論法として形式化できるものである [3]。

前提 1	雨が降ったら，地面が濡れる	( $A \rightarrow B$ )
前提 2	地面が濡れている	( $B$ )
結論	雨が降ったのだろう	( $A$ )

正しい（妥当な）三段論法とは， $A$  や  $B$  といった命題内容にかかわらず正当化可能な推論形式のことである。その意味で上記の三段論法は誤りである。なぜならその推論は， $A$  や  $B$  の内容いかんによって，真ともなれば偽ともなるからである。

実はこの推論形式は，アメリカの哲学者チャールズ・サンダース・パースが「仮説形成（アブダクション）abduction」と呼んだ，演繹 deduction と帰納 induction と異なる第 3 の推論形式にも一致している [18]。パースの所説によればアブダクションとは，人間が手さぐりの状況でわからないことをわかるための，正確さを欠くにせよ，発見的な方法である。そして，そこにはわれわれの言う対称性バイアスが含まれている。 $A \rightarrow B$  が  $B \rightarrow A$  をも同時に喚起するのであれば，上の誤った三段論法は妥当性を回復するのだからである。仮説形成的推論とは，いわば，「その可能性に賭ける」推論方法であると考えられる。

ここで重要なことは，対称性バイアスや相互排他性バイアスは，それが一種の「賭け」である以上，100% 効かなくても困るし，100% 効いていても困るということである。それらが結果的に好ましい効果を与えうるためには，それらが適度に効きつつ，しかも状況に応じて調整されるようであればならない。

これらのことから，両バイアスを重ね合わせ，これらが同時に効いたり，せめぎあったりするような推論のありようを構想することで，単純に機械的ではなく，より柔軟でより人間的なアルゴリズムを構築できるのではないかと考えられるのである。

## 2 因果関係についての期待形成モデル

人間は，原因となる事象  $p$  と結果となる事象  $q$  とのあいだにどのような因果関係を認めるのだろうか。服部 [5, 7, 8] は，それらの共起情報から因果関係を帰納することを「因果帰納 causal induction」と呼び，実験を通じてその確率論的モデルを提案している。

まず  $p$  を「ミルクを飲む」，また  $q$  を「腹痛」とする。表 1 で  $a, b, c, d$  は，おのおの  $pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}$  の共起頻度を表す。

そのうえで，被験者に  $a, b, c, d$  について具体的な数字を与える。被験者は，与えられた情報をもとに， $p$  と  $q$  のあいだの因果関係の強度を 100 段階で評価する。服部は，この被験者実験の結果得られたデータを説明しうる因果帰納のモデルとして，「随伴性モデル」などとの比較を通じて，「2 要因ヒューリスティックス・モデ

	$p$ (ミルクを飲む)	$\bar{p}$ (飲まない)
$q$ (腹痛)	$a$	$c$
$\bar{q}$ (腹痛なし)	$b$	$d$

表 1  $p, \bar{p}$  と  $q, \bar{q}$  のあいだの共起頻度

ル」の妥当性を主張している。

われわれの考えでは、この「因果帰納」によって形成された信念（確信度）は、過去の事象どうしの関係の把握や現在の事象の原因推定にだけでなく、将来の事象に対する期待形成にも役立つはずである。このため、帰納だけではなくパースの言う仮説形成をも含むという意味も含めて、われわれは服部の「因果帰納」概念を拡張して、先述のとおり「因果推論 causal inference」と呼び直すことにしたい。

ここでわれわれは、因果推論 / 期待形成のモデルとして、対称性バイアスおよび相互排他性バイアスを緩やかに含むモデル（「緩い対称性バイアス」と呼ぶ）を提案する。またこれとの対照のために、服部の検討した上記 2 つのモデルのほか、両バイアスを全く含まない条件確率モデルと、両バイアスが効きすぎているモデルを加えて、計 5 種類のモデルについて比較検討を行う。

1. 条件確率モデル *Conditional Probability Model* (CP モデル・ $P$  で表す):  
対称性・相互排他性と無関係なモデル
2. 随伴性モデル *Contingency Model* (C モデル・ $DP$ ):  
相互排他性がつねに満たされるモデル
3. 2 要因ヒューリスティックス・モデル *Dual-Factor Heuristics Model* (DFH モデル・ $H$ ):  
対称性がつねに満たされるモデル
4. 完全対称性モデル *Rigid Symmetry Model* (RS モデル・ $S_0$ ):  
対称性 + 相互排他性がつねに満たされるモデル
5. 緩い対称性モデル *Loose Symmetry Model* (LS モデル・ $S$ ):  
対称性 + 相互排他性がほどよく・柔軟に効くモデル

以下、それぞれのモデルの定義を与える。またモデル自体がもつ性質、被験者データとの適合性についても概観しておく。

## 2.1 条件確率 (CP) モデル

因果推論のモデルを構築するにあたって、最もシンプルで、他の多くのモデルのベースとなるのがこの CP モデルである。過去の条件付き確率をそのまま将来の事象の生起確率と見なす、というのがその基本的な考え方である。

$$P(q|p) = \frac{a}{a+b} \quad (1)$$

$$P(q|\bar{p}) = \frac{c}{c+d} \quad (2)$$

$$P(p|q) = \frac{a}{a+c} \quad (3)$$

$$P(\bar{q}|\bar{p}) = \frac{d}{c+d} \quad (4)$$

式 (1) は原因  $p$  が起こったときに結果  $q$  が生じる可能性を、式 (2) は原因  $p$  が起こらなかったときに結果  $q$  が生じる可能性を、それぞれ表す。その他の ( $P(p|q)$ ,  $P(\bar{q}|\bar{p})$  など以下) も同様に計算できる。

常識的にも  $p$  という条件のもとで  $q$  は起こるか起こらないかのどちらかであるが、式の上でも確かに  $P(q|p) + P(\bar{q}|p) = \frac{a}{a+b} + \frac{b}{a+b} = 1$  となる。以下この関係を満たすことを「排中律を満たす」と表現する。

「 $p$  ならば  $q$ 」を因果関係と捉えた場合、 $p$  が与えられたときに  $q$  が生起する可能性を  $P(q|p)$  と見なすのは、確かにある程度理に合っているとも思える。が、実際の人間の被験者による評価値 ([8] に基づく) と CP モデルの評価値  $P(q|p)$  とのあいだの相関を調べてみると、そこにはほとんど相関がないことが確認できる ( $r^2 = 0.0001$ )。したがってより人間的なモデルを構築するためには、少なくとも何らかの修正が必要である。

## 2.2 随伴性 (C) モデル

$p$  が  $q$  の原因であると言えるためには、どういう条件が必要なのか。おそらく、 $P(q|p)$  が高いだけでは不十分であり、 $P(q|\bar{p})$  ( $p$  が起きなかったのに  $q$  が起きる可能性) が低い、ということも、心理的には要求されるように思われる。 $p$  でも  $\bar{p}$  でも  $q$  が引き起こされるのなら、 $p$  が  $q$  の原因であるとは言えなくなるためである\*2。

因果推論の古典的なモデルとしてジェンキンスら [10] が提案するところによれば、そのモデルとなる式は、 $P(q|p)$  の高さに加えて  $P(q|\bar{p})$  の低さを考慮して、

$$\Delta P(q|p) = P(q|p) - P(q|\bar{p}) \quad (5)$$

$$\Delta P(q|\bar{p}) = P(q|\bar{p}) - P(q|p) \quad (6)$$

などと書けるといふ。

ここでわれわれのバイアスの観点から言えば、(5) 式の  $p$  と  $q$  の場所におののおの  $\bar{p}$  と  $\bar{q}$  を代入すると、

$$\begin{aligned} \Delta P(\bar{q}|\bar{p}) &= P(\bar{q}|\bar{p}) - P(\bar{q}|p) \\ &= P(\bar{q}|\bar{p}) - (1 - P(q|p)) \\ &= \Delta P(q|p) \end{aligned}$$

となり、この式が相互排他性をつねに満たすことが確認できる。ただし対称性は満足しない。

なおこの場合、 $\Delta P$  は  $[-1, 1]$  区間の値をとるが、比較の都合上  $[0, 1]$  区間の値をとるように線形変換を施し、変換後  $DP$  と書くこととする。

$$DP(q|p) = \frac{P(q|p) - P(q|\bar{p}) + 1}{2} \quad (7)$$

$$DP(q|\bar{p}) = \frac{P(q|\bar{p}) - P(q|p) + 1}{2} \quad (8)$$

このとき、 $DP(q|p) + DP(q|\bar{p}) = 1$  となり、排中律を満足する。

$DP$  を  $a, b, c, d$  を用いて書き直せば、以下のような式となる。

$$DP(q|p) = \frac{1}{2} \left( \frac{ad - bc}{(a+b)(c+d)} + 1 \right) \quad (9)$$

$$DP(q|\bar{p}) = \frac{1}{2} \left( \frac{bc - ad}{(a+b)(c+d)} + 1 \right) \quad (10)$$

\*2 このこと自体が、相互排他性バイアスが存在するということの別表現でもある。「 $p$  が起きるならば  $q$  も起きる可能性が高い」ということは、「 $p$  が起こらなければ  $q$  も起こらない可能性が高い」ということである気がする。すなわち  $(p \rightarrow q) \implies (\bar{p} \rightarrow \bar{q})$  である。

被験者による評価値と C モデルの評価値  $DP(q|p)$  とのあいだの相関も、やはりほとんど見られなかった ( $r^2 = 0.0$ )。

## 2.3 2 要因ヒューリスティックス (DFH)・モデル

服部が因果推論のモデルとして提唱する DFH モデルのエージェントは、以下の式を用いて期待を形成する。

$$H(q|p) = \sqrt{P(q|p)P(p|q)} \quad (11)$$

$$H(q|\bar{p}) = \sqrt{P(q|\bar{p})P(\bar{p}|q)} \quad (12)$$

式から明らかなように、 $H$  は対称性 (服部は「双条件性 biconditionality」と呼ぶ [5]) をつねに満足する ( $H$  は  $P(q|p)$  とその対称式の相乗平均である)。しかし排中律は必ずしも満たさず ( $H(q|p) + H(\bar{q}|p) \neq 1$ )、相互排他性も満たさない。

$P(q|p) = \frac{a}{a+b}$ 、 $P(p|q) = \frac{a}{a+c}$  より共起頻度  $a, b, c, d$  で書き直せば、

$$H(q|p) = \frac{a}{\sqrt{(a+b)(a+c)}} \quad (13)$$

$$H(q|\bar{p}) = \frac{c}{\sqrt{(a+c)(c+d)}} \quad (14)$$

となる。ここで注目すべきは、 $H(q|p)$  の計算には値  $d$  が、 $H(q|\bar{p})$  の計算には値  $b$  が使用されていないということである。より少ない情報で推測を行うヒューリスティックスたるゆえんである。

このモデルの長所は、このモデルの出力値が人間による評価値ときわめて強い相関 ( $r^2 = 0.96$ ) をもつことである [8]。

## 2.4 完全対称性 (RS) モデル

他のモデル、特に後述の LS モデルとの比較対照のため、排中律、対称性、相互排他性のすべてが つねに成立するモデルを考える。これはきわめて強い対称性をもつため、完全対称性モデルと呼ぶことにする。

なお、上記 3 つの条件すべてを同時に満たす式はこれ以外にも存在するので、この定義が唯一というわけではない。

$$S_0(q|p) = \frac{a+d}{a+b+c+d} \quad (15)$$

$$S_0(q|\bar{p}) = \frac{b+c}{a+b+c+d} \quad (16)$$

$$S_0(p|q) = \frac{a+d}{a+b+c+d} \quad (17)$$

$$S_0(\bar{q}|\bar{p}) = \frac{a+d}{a+b+c+d} \quad (18)$$

これについても人間の評価値との相関を調べたが、弱い相関しか見られなかった ( $r^2 = 0.158$ )。やはり人間の推論にこれほどまでに強く両バイアスがかかっているということは考えにくい。

## 2.5 緩い対称性 (LS) モデル

われわれの本命はこの LS モデルである。第 1.3 節で考察したように、われわれは、対称性と相互排他性の両方を緩やかに・かつ柔軟に含むモデルが高いパフォーマンスを出せるはずだと予想した。そして試行錯誤の

結果，次のような式がその狙いを満たすものであることを見いだした。

$$S(q|p) = \frac{a + \left(\frac{b}{b+d}\right)d}{a + b + \left(\frac{a}{a+c}\right)c + \left(\frac{b}{b+d}\right)d} \quad (19)$$

$$S(q|\bar{p}) = \frac{\left(\frac{d}{b+d}\right)b + c}{\left(\frac{c}{a+c}\right)a + \left(\frac{d}{b+d}\right)b + c + d} \quad (20)$$

$$S(p|q) = \frac{a + \left(\frac{c}{c+d}\right)d}{a + \left(\frac{a}{a+b}\right)b + c + \left(\frac{c}{c+d}\right)d} \quad (21)$$

$$S(\bar{q}|\bar{p}) = \frac{\left(\frac{c}{a+c}\right)a + d}{\left(\frac{c}{a+c}\right)a + \left(\frac{d}{b+d}\right)b + c + d} \quad (22)$$

かなり複雑に見えるが，特に (19) 式のみに着目すると，それは

$$S(q|p) = \frac{a + \beta d}{a + b + \alpha c + \beta d} \quad (23)$$

$$\text{ただし } \alpha = \frac{a}{a+c}, \quad \beta = \frac{b}{b+d} \quad (24)$$

という形式になっている。ここでもし  $\alpha = \beta = 0$  であったならばこの式は  $\frac{a}{a+b}$  となり， $P(q|p)$  に一致する。また  $\alpha = \beta = 1$  であったならば  $\frac{a+d}{a+b+c+d}$  となり， $S_0(q|p)$  に一致する。したがって，この  $S$  なる式こそが一般形であり， $P$  と  $S_0$  が両極端をなす特殊形であるとも考えることもできる。つまり LS モデルは，両バイアスと無関係な CP と，両バイアスが効きすぎている RS との中間的な性質をもつと考えられる。

モデル内部での対称性バイアスの効き具合を調べるには  $P(q|p)$  と  $P(p|q)$  の相関を，相互排他性バイアスの効き具合を調べるには  $P(q|p)$  と  $P(\bar{q}|\bar{p})$  との相関を見ればよい。図 2 および 3 において，LS が CP (グラフは一様分布になる) と RS (グラフは対角線になる) の中間であることが視覚的に確認できる。

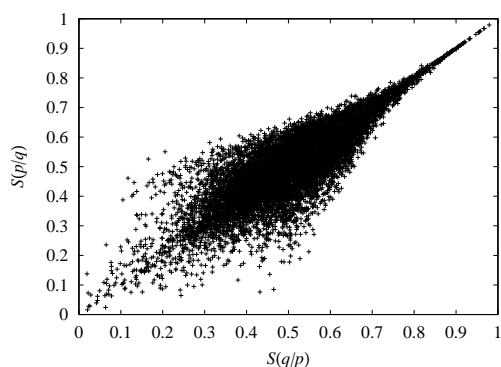


図 2 LS モデルにおける対称性バイアス

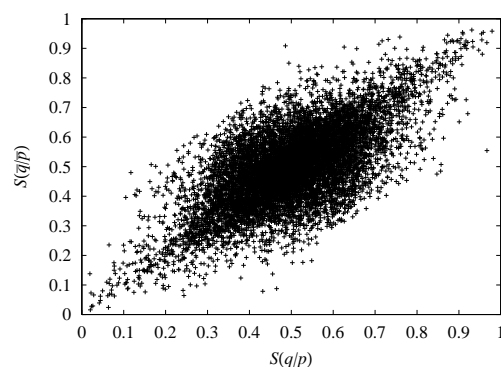


図 3 LS モデルにおける相互排他性バイアス

またこの LS モデルの特徴は，DFH モデル以上に人間の評価値（人間の感覚）と合致する点である ( $r^2 = 0.97$ )。ちなみに，上記  $\alpha$  と  $\beta$  をどのような定数にしても人間の評価値とは合致しなかった。

このように，人間の因果の感覚と合致するかしないかについて，さしあたり相関関係のみを根拠として言えば，これらのモデルの中で「人間らしい」のは DFH と LS であると言える。

### 3 実験の設定と結果

#### 3.1 2本腕バンディット問題

前節まででわれわれは、LSのような定義と性質をもつ推論モデルがさまざまな課題に対して人間に似た・人間並のパフォーマンスを示すはずだ、という仮説を立てた。この仮説の検証のため、本報告では強化学習などの分野で典型的な課題としての「2本腕バンディット問題 2-armed Bandit Problem」を試金石とする。

一般に、「 $N$ 本腕バンディット問題  $N$ -armed/Multi-armed Bandit Problem」の構造を理解するには、 $N$ 本の腕のあるスロットマシンをイメージするとよい。それぞれの腕の当たり確率は事前に決められており、腕を引けば、その確率にしたがってコイン（報酬）が出てくる。問題は、プレイヤーにとってどの腕が有利なのかが事前にはわからないことである。彼/女は、最初はさまざまな腕を試してみて、どの腕が有利なのかをできるだけ早く判断し、あとはその有利な腕ばかり引く（「貪欲 greedy 法」）ことで報酬を最大化したいところである。しかしながら、どの腕が有利なのかを正確に判断するためには、より多く試してみることで、つまり「探索」が必要である。したがってプレイヤーが報酬を最大にするためには、この問題に内在するとされている「探索 exploration と知識利用 exploitation のトレードオフ」[23]、あるいは言い換えれば「正確さと早さのトレードオフ」を無理やり克服しなければならないことになる。

この問題は、教科書に載っているような単純で典型的なものではあるが、人間の意思決定や、経験的学習の際の困難を本質的に捉えていると考えられる。たとえば、旅先など見知らぬ場所でレストランに食事に行く際、どの店に入るべきかを判断するにはどうしたらよいただろうか。ガイドブックや口コミのような鳥瞰的な情報が得られない場合には、ともかく目ぼしい店に実際に入って食べてみるしかない。そのようにして「探索」を繰り返せば、経験的にどの店が自分にとって満足度の高い店なのかが徐々にわかってゆくことになる。

あるいは、セールスマンが近隣の街に売り込みに行くときに、すべての街を回るのは物理的に困難であるので、どの街に行けばより売れるのかを予測し、そこを重点的に売り込みに行くという場合を考えれば、これもやはり  $N$ 本腕バンディット問題の形式をもつことになる。実際に行き、売ってみるまでは何もわからない。彼/女は実際に売り歩いてみて、得られた経験から判断し、よく売れる街を優先的に回るようにするであろう。

このように、「正確に判断したいが、ともかくやってみなければわからない」という状況は日常生活にあふれており、かつ、そこには「探索と知識利用のトレードオフ」が同様に隠れている。この意味で、 $N$ 本腕バンディット問題は、意思決定や学習に際する人間の一種ゲーム論的ジレンマ状況を縮約するよう思われるのである。

われわれはすでに、(1) いかにか「人間らしく」振る舞うか、という定性的な観点からモデルの特性を比較した。ここではさらに、(2) この問題にいかにか多くの報酬を得るかという定量的観点から実験を行い、結果を分析した。

#### 3.2 課題設定に関するパラメーター

まず課題の設定に関して、着眼点となる重要なパラメーターについて定義しておく。なお、シミュレーションはランダム・シードを変えて 100,000 回行った。

腕の本数 今回は最も単純なケースに絞るため、 $N = 2$  とする。

各腕の客観的報酬確率 あらかじめ腕 A に定められている報酬確率を  $P_A = \hat{P}(q|p)$  , 腕 B のそれを  $P_B = \hat{P}(q|\bar{p})$  とする。ただし  $0 \leq P_A, P_B \leq 1$  とする。

難易度・難易度レベル 2 本腕バンディット問題には易しい問題と難しい問題がありうる。2 本の腕の報酬確率が大きく異なる場合、どちらが有利な腕なのかを判断するのは難しくない。他方、両確率が接近していれば、判断は難しくなる。

そこで、問題の「難易度 difficulty」というパラメーターを以下のように定義する。問題は、難易度が 0 に近いほど易しくなり、1 に近いほど難しくなると考えられる。

$$(\text{難易度}) = 1 - |\hat{P}(q|p) - \hat{P}(q|\bar{p})|$$

また、この難易度のスケールを 10 段階に均等に分割し、「難易度レベル difficulty level 1 ~ 10」と呼ぶ。

正解率 シミュレーション 100,000 試行のうち、「正しい」(客観的報酬確率の高いほうの)腕を引いた割合を「正解率 correct rate」と呼ぶ。

与えられた条件下で報酬の最大化を求めるということは、究極的にはこの正解率を求めると一致する。

探索率 シミュレーション期間内に選択肢を変更した割合を「探索率 exploration rate」と呼ぶ。これも 100,000 回の平均値である。

修正条件確率モデル 2 本腕バンディット問題の特性である正確さと早さのトレードオフ関係を視覚化するために、CP モデルのアドホックな修正モデル (MCP モデル) を考える。

CP モデルは「貪欲法」、すなわち、ひとたび有利なほうと判断したらその腕ばかりを引く方法をとるが、ここではその修正版である「 $\epsilon$ -貪欲法」をさらに修正した方法を用いる。

$\epsilon$ -貪欲法とは、基本的には貪欲法だが  $\epsilon$  の確率だけ腕をランダムに選ぶ方法である。通常はこの  $\epsilon$  を一定とするが、この上、時間が経つにつれて探索の必要は減少するはずである。

そこで、ここでは  $\epsilon$  の値を時間とともに減少させるパラメーター  $\tau$  を導入し、

$$\epsilon = \frac{0.5}{1.0 + \tau t} \quad (25)$$

という式にしたがって  $\epsilon$  を減衰させる。  $\tau$  が大きければ大きいほど  $\epsilon$  は急速に減少する。  $\tau = \infty$  の場合  $\epsilon = 0$  となり、MCP は CP と一致する。

$\tau$  は 0.05 ~ 1 のあいだで 0.05 刻みで与えた。

### 3.3 実験結果とその評価

シミュレーションの結果、報酬の獲得という観点からは、われわれの予想を超えるほどに、LS モデルがよい成績を収めた。まず LS モデルは他のモデルに比べ、獲得報酬において、すなわち正解率において圧倒的に高い値を記録している (図 4)。

また、図 5 からは、CP モデルと MCP モデルが先述の「探索と知識利用のトレードオフ」ないし「正確さ (縦軸) と早さ (横軸) のトレードオフ」に頑強に拘束されていることが見て取れる (右下がりの点線)。これに対し、LS モデルは易々とこのトレードオフを乗り越え、早さにおいても正確さにおいても高い水準を両立している。他方 DFH モデルは、残念ながらこの課題に対しては CP よりもかなりパフォーマンスが悪い。

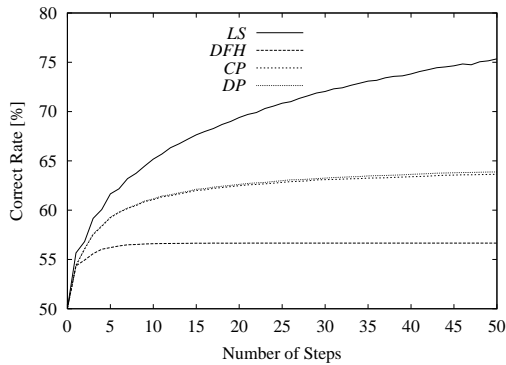


図4 試行回数と正解率

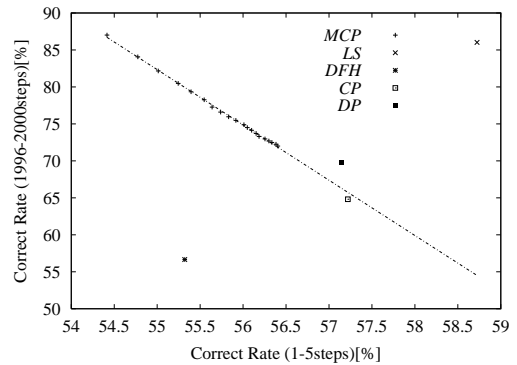


図5 正確さと早さのトレードオフ

予想されたこととはいえ、LSモデルがなぜこのように高いパフォーマンスを示すのか。課題の難易度に着目してその理由を探ると、他のモデルとは全く異なり、LSモデルは難易度が上がるにつれ（半分＝中央値を超えると）探索率を自律的に急激に上げているのが観察できた（図6）。

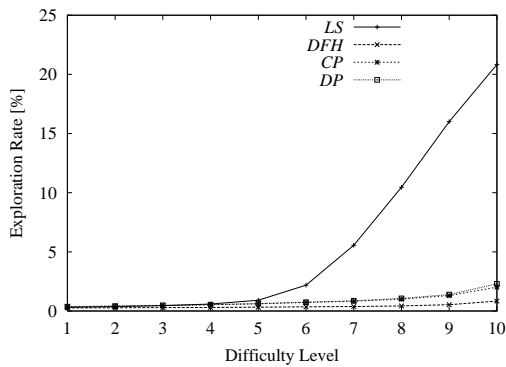


図6 難易度に応じた探索率の自律的調整

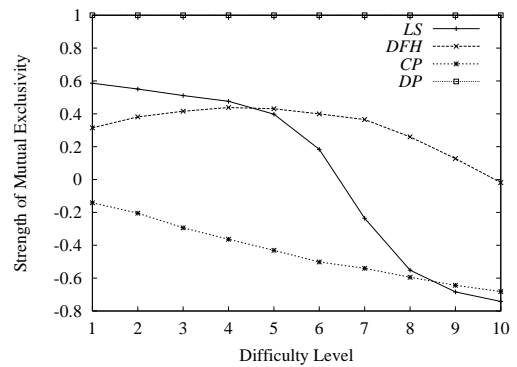


図7 難易度に応じた相互排他性強度の変化

このとき内部状態を見てみると、難易度が中央値を超えると、相互排他性バイアスの効き具合が負の値を示していることがわかった（図7）。すなわち問題自体が一定以上の難易度をもつ場合、LSの挙動に逆相互排他性バイアスがかかるということである。

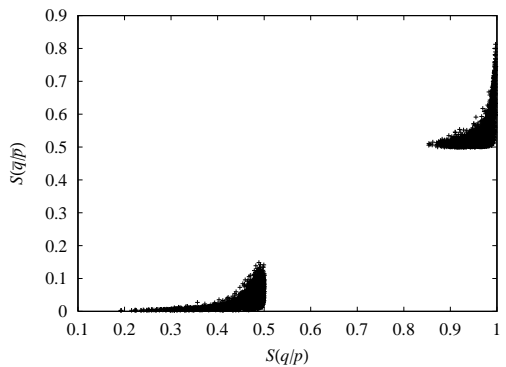


図8 難易度レベル1における相互排他性

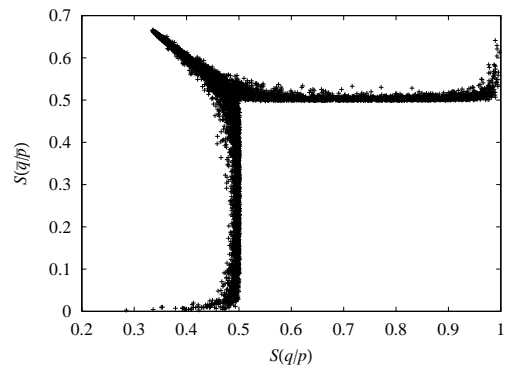


図9 難易度レベル10における相互排他性

負の相互排他性バイアスが効くということは、たとえば  $p \rightarrow q$  の可能性が高いという情報から  $\bar{p} \rightarrow \bar{q}$  の可能性が低い ( $\bar{p} \rightarrow q$  の可能性が高い) という期待を形成する傾向を生じるということに等しい。すなわち、腕 A を引くことで報酬獲得の頻度は経験上多かったけれども、素直にそれを信じず、実は腕 B を引いた方がもっとよいのではないか、という気持ちをもつことである。すなわち、LS モデルのエージェントは問題が難しくなると「迷う」「悩む」ようになるのである。

「いま引いている腕は、有利なほうの腕ではないのではないか。もう一方の腕のほうが、実は有利なのではないか」という疑念をもつことによって探索率が高まる。難しい問題においてはその定義上どちらの腕を引いても報酬差が少ないので探索によって失うものが少なく、より迷うことで正確さにおいて優位となるだろう。この結果正解率が高まり、試行全体として多くの報酬を得ることができるのだと考えられる。

ここでこの課題の中でエージェントに課されていることの意義を考えてみよう。先にも述べたように、報酬の最大化を求めるということは、正解率を求めることである。つまりエージェントに課されているのはもはや両腕の報酬確率を計算し比較することであるよりは、端的に より有利な腕がどちらなのかを予想することであり、絶対評価ではなく相対評価で期待を形成することである。LS モデルでは、たとえば腕 A がこれまでのところ確率がよかったとしても、腕 B のほうにもまだ可能性を認めるということが可能である。このことによって、難易度と無関係に一定の戦略を貫く他のモデルには超えられないトレードオフが超えられるのである。

LS モデルのこうした挙動は、モデル自体が定義上決定論的なものであるにもかかわらず、機械的な硬直性を超えた、それ自体人間らしい柔軟性であるように見える。

## 4 結 論

この問題を定量的な意味で解くということだけが目的ならば、いくらでもアドホックな修正を施すことができ、獲得報酬の最大化は図れるであろう。しかし本報告の趣旨は、より人間らしいアルゴリズムを探究することであって、特定種類の問題に強いアルゴリズムを探すことではない。

この LS モデルは、かなりエレガントなかたちのままで、すなわち決定論的なモデルとして、2 本腕バンディット問題に対して十分な好成績を出すことができた。また同時に、2 つの意味で「人間的な」振る舞いを示した。一つはモデル自体が「人間の感覚」と強い相関をもつという意味で、もう一つは悩ましい問題に際しては「迷う」という挙動、言い換えれば課題の難易度によって戦略を変えるという柔軟性を示すという意味である。

以上のことから、対称性バイアスと相互排他性バイアスという 2 つの非論理的バイアスを緩く含んだ LS モデルは、人間の「論理的ではない」推論・意思決定モデルとして非常に優秀なものの一つであると結論づけてよいと思われる。もちろん人間が実際にこの複雑な式にしたがって推論や意思決定に際しその都度計算をしているとは考えられないが、このモデルがこの課題における人間の推論を結果的によく表現していることは確かであろう。

ただし現在のところ、検証は 2 本腕バンディット問題とその他若干の課題 [20] についてしかなされていない。本報告の課題の構造が偶然このモデルには向いていた、という可能性も否定はできない。モデルの一般的な有効性は、今後さまざまな問題を課すことで検証されてゆく必要があるだろう。

しかし少なくとも本報告において、対称性バイアスおよび相互排他性バイアスという 2 つのバイアスに関しては、使いようによってはそれらが非常に有効であること、また、これらを独立に扱うよりも組みあわせて使うことが有効であること、この 2 点については立証できたと考える。人間の思考・推論・学習・意思決定を探究するあらゆる分野において、この結論が今後支持されることが期待される。

## 参考文献

- [1] Davidson, D. and Tell, D. 2005 “Monolingual and Bilingual children’s Use of Mutual Exclusivity in the Naming of Whole Objects,” *Journal of Experimental Child Psychology*, Vol. 92, pp. 25–45.
- [2] Doyle, Sir Arthur Conan 1887 *A Study in Scarlet*. 『緋色の研究』.
- [3] Eco, U. and Sebeok, T. A. (eds.) 1988 *The Sign of Three: Dupin, Holmes, Peirce*, Indiana University Press. エーコ/シービオク編 1990 『三人の記号 デュパン/ホームズ/パース』, 東京図書.
- [4] Halberda, J. 2003 “The Development of a Wordlearning Strategy,” *Cognition*, Vol. 87, pp. B23–B34.
- [5] 服部 雅史 2001 「因果帰納の二要因ヒューリスティックス・モデル」, 『認知科学』, Vol. 8, pp. 444–453 .
- [6] 服部 雅史 2002a 「条件文推論における方向性」, 『立命館人間科学研究』, Vol. 3, pp. 1–13.
- [7] 服部 雅史 2002b 「論理的推論への確率論的アプローチ」, 『学習と対話』, Vol. 1, pp. 16–26.
- [8] Hattori, M. 2003 “Adaptive Heuristics of Covariation Detection: A Model of Causal Induction,” in *Proceedings of the 4th International Conference on Cognitive Science and the 7th Australasian Society for Cognitive Science Joint Conference (ICCS/ASCS 2003)*, Vol. 1, pp. 163–168.
- [9] Imai, M., Gentner, D., and Uchida, N. 1994 “Children’s Theory of Word Meanings: The role of shape similarity in early acquisition,” *Cognitive Development*, Vol. 9, pp. 45–75.
- [10] Jenkins, H. M. and Ward, W. C. 1965 “Judgment of contingency between responses and outcomes,” *Psychological Monographs*, Vol. 79, pp. 1–17.
- [11] 岸田 秀 1977(1996) 『ものぐさ精神分析』, 中公文庫 .
- [12] Markman, E. M. and Wachtel, G. F. 1988 “Children’s Use of Mutual Exclusivity to Constrain the Meaning of Words,” *Cognitive Psychology*, Vol. 20, pp. 121–157.
- [13] Markman, E. M. 1989 *Categorization and Naming in Children: Problems of induction*, MIT Press, Cambridge.
- [14] Markman, E. M. 1990 “Constraints Children Place on Word Meanings,” *Cognitive Science*, Vol. 14, pp. 121–157.
- [15] Markman, E. M., Wasow, J. L., and Hansen, M. B. 2003 “Use of the Mutual Exclusivity Assumption by Young Word Learners,” *Cognitive Psychology*, Vol. 47, pp. 241–275.
- [16] J・マッカーシー, 松原 仁, P・J・ヘイズ 1990 『人工知能になぜ哲学が必要か フレーム問題の発端と展開』, 哲学書房.
- [17] 中野 昌宏, 篠原 修二 2006 「コミュニケーションにおける Aha! は, 『話が裏返る』 ことによってもたらされる」, 『InterCommunication』, No. 58, NTT 出版, pp. 53–65 .
- [18] Peirce, C. S. 1898(1992) *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, ed. Kenneth Laine Ketner, Harvard University Press. C・S・パース 2001 『連続性の哲学』, 伊藤邦武訳, 岩波文庫.
- [19] 篠原 修二, 田口 亮, 桂田 浩一, 新田 恒雄 2007 「因果性に基づく信念形成モデルと  $N$  本腕バンディット問題への適用」, 『人工知能学会論文誌』, Vol. 22, No. 1, pp. 58–68.
- [20] 篠原 修二, 田口 亮, 桂田 浩一, 新田 恒雄 2007 「語彙学習エージェントにおけるバイアスの自律調整について」, 『人工知能学会論文誌』, Vol. 22, No. 2, 掲載予定.
- [21] 塩沢 由典 1990(1998) 『市場の秩序学 反均衡から複雑系へ』, 筑摩書房.

- [22] Sidman, M. et al. 1982 “A Search for Symmetry in the Conditional Discriminations of Rhesus Monkeys, Baboons, and Children,” *Journal of the Experimental Analysis of Behaviour*, Vol. 37, pp. 23–44.
- [23] Sutton, R. S. and Barto, A. G. 1998 *Reinforcement Learning: An Introduction*, MIT Press, Cambridge: MA.
- [24] Wason, P. C. 1966 “Reasoning” in Foss, B. M. (ed.), *New Horizons in Psychology*, Penguin Books, Harmondsworth: Middlesex, UK, pp. 135–151.
- [25] 山崎 由美子 1999 「動物における刺激等価性」, 『動物心理学研究』, Vol. 49, pp. 107–137.
- [26] Yamazaki, Y. 2004 “Logical and illogical behavior in animals,” *Japanese Psychological Research*, Vol. 46, pp. 195–206.
- [27] 山崎 由美子, 日原 さやか, 藤井 直敬, 岡ノ谷 一夫, 入来 篤史 2006 「概念の発達と操作の神経機構 ヒト思考形式の非論理バイアスによる概念創発」, 『生体の科学』, Vol. 57, pp. 51–57.
- [28] 山崎 由美子, 岡ノ谷 一夫, 入来 篤史 (2006) 「動物の論理を比較する 意味の獲得へと導くメカニズム」, 『信学技報』, Vol. NC2005-93 (2006-1), pp. 19–24.