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

A Model of Belief Formation Based on Causality and Application to N -armed Bandit Problem

篠原 修二

Shuji Shinohara

豊橋技術科学大学大学院工学研究科知識情報工学系

Department of Knowledge-based Information Engineering, Graduate School of Engineering, Toyohashi Univ. of Tech.
shinohara@vox.tutkie.tut.ac.jp

田口 亮

Ryo Taguchi

(同 上)

taguchi@vox.tutkie.tut.ac.jp

桂田 浩一

Kouichi Katsurada

(同 上)

katurada@tutkie.tut.ac.jp

新田 恒雄

Tsuneo Nitta

(同 上)

nitta@tutkie.tut.ac.jp

keywords: causal induction, symmetry bias, mutual exclusivity bias, n -armed bandit problem, trade-off between exploration and exploitation

Summary

Through numbers of studies on the formation of equivalence relations and causal induction, it is known that human beings tend to consider conditional statements “if p then q ” as biconditional statements “if and only if p then q ”: we call the tendency to perceive “if p then q ” as “if q then p ” the “symmetry bias”. On the other hand, many studies on children’s word learning have pointed out that children tend to expect each object has only one label. This is so-called the “mutual exclusivity bias”. This bias implies that children infer “if not p then not q ” from “if p then q ”.

These biases logically mislead human beings. What is the merit of these illogical induction? In this paper we address this question. First, we clarify the relationship between causal induction and the symmetry bias or and the mutual exclusivity bias. Secondly, we propose a new model of causal induction. Thirdly, we construct an agent which makes illogical decision based on causality, and assess the agent’s performance for the task called “ n -armed bandit problem” in the field of reinforcement learning. In this problem, it is known that there is a “trade-off between exploration and exploitation”. According to our simulation, the agent can resolve the trade-off and achieve quite better performance than an agent without these biases.

1. はじめに

刺激等価性の成立や因果帰納に関する研究を通して、ヒトは「 p ならば q 」から「 q ならば p 」を推論する傾向性を持つことが知られている [Sidman 82, 山崎 06a, 山崎 99, Yamazaki 04, 山崎 06b, 服部 02a]。身近な例を挙げれば、我々は「もし部屋を片付けたら、遊びに連れて行ってあげる」から「もし部屋を片付けたら、その時だけ遊びに連れて行ってあげる」を推論してしまう [服部 02a]。このような推論は、論理的に正しいとはいえないにもかかわらず、年齢や能力などとは無関係に多くのヒトで容易に成立する。一方、ヒト以外の動物にとってはこのような対称性の成立は困難であることが知られており、ヒトとそれ以外の動物でなぜこのような違いを示すのかが注目されている [田中 00, 藤田 01]。以後本稿で

は「 p ならば q 」から「 q ならば p 」を推論する傾向性のことに対称性バイアスと呼ぶことにする。

一方、幼児の言語獲得に関する研究において、幼児は様々なバイアスを持つことで、効率的に語彙学習を行っているという考えが提唱されている [Markman 89, Imai 94]。語彙学習バイアスとして様々なものが提案されているが、その一つに相互排他性バイアスがある。これは、対象は一つのラベルのみを持つと考える傾向性のことである [Markman 88, Markman 90, Markman 03, Halberda 03, Davidson 05]。例えば既知の事物と新奇な事物を目の前にして新奇なラベルを聞いた時、幼児はそのラベルを新奇な事物の方の名前であるとみなす。このバイアスは指差しなどが無い場合に指示対象の同定に役立つと考えられている [Markman 03]。ここで既知のラベルと既知の対象を、おのおの l_{known} , o_{known} と表記すること

にしよう。また、ある対象が l_{known} と呼ばれることを p , o_{known} であることを q と表記するとすれば、幼児は、「 p ならば q 」ということを知っている時、「 p でないならば q でない」と信じる傾向性を持つといえる^{*1}。これは冒頭で述べた例を用いるなら、「もし部屋を片付けたら、遊びに連れて行ってあげる」から「もし部屋を片付けないなら、遊びに連れて行ってあげない」を推論してしまうということである。対称性バイアスの場合と同様、「 p ならば q 」から「 p でないならば q でない」を推論することは、必ずしも真ではない。

以上に示した傾向性は、人間に固有なものである一方、論理的に誤った結論を導く。論理的に間違っているにもかかわらず、このようなバイアスを適用することのメリットは何なのだろうか。換言すれば、このような非論理的推論が有効に作用する場面とはどのような状況なのだろうか。本稿では、この問題を扱う。

2 章では、因果関係の帰納に関する確率モデルをいくつか紹介した後、これらのモデルと対称性バイアス、相互排他性バイアスとの関連について述べる。3 章では、バイアスを組み込んだ因果関係の帰納に関するモデルを提案する。4 章では、非論理的推論が有効に作用することが期待される問題として、強化学習の分野で「 N 本腕バンディット問題」と呼ばれている課題を取り上げる^{*2}。ただし本稿では特に「2 本腕バンディット問題」を扱う。5 章では、提案モデルを 2 本腕バンディット問題に適用したシミュレーション結果を示す。6 章はまとめとなっている。

2. 因果関係の帰納に関するモデル

服部は、因果帰納に関する研究を通して、人間は「 p ならば q 」という条件文を双条件的に解釈する傾向性を持つとともに、双条件性が守られている時に因果関係が認知されやすいと指摘する [服部 01, 服部 02b, 服部 02a]。また Goswami は、幼児は生まれながらにして因果に注目する傾向を備えており、この傾向が認知発達の最も基本的なプロセスであると主張する [Goswami 98]。例えば「因果推論は、発達のはじまりから、物理的世界の出来事を解釈し、表象し、記憶するために重要な役割を果たしている」と述べている。では事象間の関係を因果関係として捉えるとはどのようなことなのだろうか。

Hattori は、原因の候補となる事象 p と結果事象 q の共起情報からどのような因果関係が帰納されるのかについての確率モデルとして、以下に示すような二要因ヒューリスティックス・モデル (dual-factor heuristics model;

以下では DH モデルと呼ぶ) を提案している^{*3} [Hattori 03]。

$$H(q|p) = \sqrt{P(q|p)P(p|q)} \quad (1)$$

ただし、 $P(q|p)$ と $P(p|q)$ はおののお、 p が起きた時に q が起きた条件付確率、 q が起きた時に p が起きていた条件付確率を表す。

DH モデル以外にも、因果関係の帰納に関して以下のようなモデルが提案されている [Jenkins 65]。

$$\Delta P(q|p) = P(q|p) - P(q|\bar{p}) \quad (2)$$

ただし、 $\bar{p}$ は p の否定、すなわち事象 p が起きなかったことを表す。

ある事象 p が起きた時に、別の事象 q が起きる確率が高い場合、 p は q の生起の予測に役立つ。この意味で、両モデルの右辺第一項に現れる条件付確率 $P(q|p)$ は、予測可能性を表しているといえる [服部 01]。もちろん因果関係の強度は、予測可能性のみによって規定されるものではない。我々の日常感覚から言えば、 p が q の原因であるといえるためには、 p が起きた時に q が起きる確率、すなわち $P(q|p)$ が高いだけではなく、同時に以下の条件も要請されるように思われる。

- (1) q が起きた時に p が起きていなかった確率 $P(\bar{p}|q)$ が低い
- (2) p が起きなかった時に q が起きない確率 $P(\bar{q}|\bar{p})$ が高い

前者は「原因のないところに結果は起こらない (火の無いところに煙は立たない)」という信念を確率的に表現したものである。後者は、 p が起きた時に q が起きる確率が高いとしても、 p が起こらなかった時にも q が高い確率で起こる、つまり p とは無関係に q が高い確率で起こるならば、 p が q の原因であるとはいえないということを表す。

DH モデルは、

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

と変形できる。また ΔP モデルは、

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

と書き換えることができる。つまり、共起情報から因果関係を帰納する際、DH モデルは予測可能性に加えて条件 (1) を、一方 ΔP モデルは条件 (2) を考慮したものと見える。

*1 対称性と相互排他性はともに生後 17 ヶ月頃に獲得されるという [Halberda 03, 山崎 06b]。もしかすると両者の成立には何らかの関連があるのかもしれない。

*2 本稿の目的は、「 N 本腕バンディット問題」を題材として、人間に見られる非論理的推論の有効性を示すことであり、この課題に対する最適アルゴリズムの提案ではないことを断っておく。

*3 [Hattori 03] において (1) 式の左辺は H と表記されているが、本稿では、前件と後件を明確にするため、 $H(q|p)$ と表記することにする。これは以下に述べる ΔP に関しても同様である。

ここで因果律に関するこれら二つのモデルを対称性バイアスと相互排他性バイアスの観点から眺めてみよう。まず、DH モデルに関して、(1) 式の p と q を入れ替えてみれば

$$\begin{aligned} H(p|q) &= \sqrt{P(p|q)P(q|p)} \\ &= H(q|p) \end{aligned}$$

となり、対称性が成立する。一方 ΔP モデルに関して、(2) 式の p と q におのおの $\bar{p}$ と $\bar{q}$ を代入すれば、

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

となり、相互排他性が成立することがわかる。つまり確率モデルの観点からいえば、「 p ならば q 」を因果関係として捉えるということは、対称性や相互排他性バイアスを持つことであるともいえる。

3. 信念形成モデル

本章では対称性と相互排他性バイアスを組み込んだ信念形成モデルを提案する。そのようなモデルを構築するにあたって、両バイアスの関係について整理しておく。

まず共起情報から帰納された「 p ならば q 」という条件文に対する主体の確信度を $B(q|p)$ (ただし $0 \leq B(q|p) \leq 1$) と表記する。以後、この確信度のことを信念と呼ぶ。条件付確率では、 $P(q|p) = 1 - P(\bar{q}|p)$ は成立する。 B に関しても同様に

$$B(q|p) = 1 - B(\bar{q}|p) \quad (3)$$

が成立するとする。これは「 p ならば q 」と信じる時、「 p ならば $\bar{q}$ 」とは信じない傾向性を持つということを表す。以後この傾向性を否定バイアスと呼ぶ。

また「 p ならば q 」を信じる時「 q ならば p 」も信じる傾向性 (対称性バイアス) を

$$B(q|p) = B(p|q) \quad (4)$$

として、「 p ならば q 」を信じる時「 $\bar{p}$ ならば $\bar{q}$ 」も信じる傾向性 (相互排他性バイアス) を

$$B(q|p) = B(\bar{q}|\bar{p}) \quad (5)$$

で表す。

ここで、主体が否定バイアスを持つという前提のもとで、対称性バイアスを持てば、相互排他性が成立することを示す。

$$\begin{aligned} B(q|p) &= 1 - B(\bar{q}|p) \text{ (否定バイアスより)} \\ &= 1 - B(p|\bar{q}) \text{ (対称性バイアスより)} \\ &= 1 - (1 - B(\bar{p}|\bar{q})) \text{ (否定バイアスより)} \\ &= B(\bar{p}|\bar{q}) \\ &= B(\bar{q}|\bar{p}) \text{ (対称性バイアスより)} \end{aligned}$$

以上の式変形から、主体が否定バイアスと対称性バイアスを持つ時、相互排他性が成立するだけでなく、 $B(q|p) = B(\bar{p}|\bar{q})$ という関係も成立することがわかる。これは「 p ならば q 」を信じる時「 $\bar{q}$ ならば $\bar{p}$ 」も信じる、すなわち対偶が成立する傾向性を表す。一方、主体が相互排他性バイアスを持つからといって対称性が成立するわけではない。つまり、否定バイアスを持つという条件のもとでは、相互排他性バイアスよりも対称性バイアスの方がより強いバイアスであるといえる。

バイアスを組み込んだモデルの候補として、前章で述べた DH モデルや ΔP モデルが挙げられる。しかし DH モデルは対称性は満足するが、否定バイアスを満足しない。このため相互排他性も成立しない。一方 ΔP モデルは、相互排他性バイアスを満足する。否定バイアスに関しては、 $\Delta P(q|p) + \Delta P(\bar{q}|p) = 0$ が成立する。 ΔP は $[-1.0, 1.0]$ 区間の値を取るが、ここで $0.0 \leq \Delta P \leq 1.0$ となるように線形変換を施す。

$$\Delta P(q|p) = \frac{P(q|p) - P(q|\bar{p}) + 1}{2} \quad (6)$$

この時、 $\Delta P(q|p) + \Delta P(\bar{q}|p) = 1$ となり否定バイアスを満足する。以後この式を ΔP モデルと呼ぶ。 ΔP をこのように再定義するならば、否定と相互排他性は成立するが、対称性バイアスは満足しない。

これら三つのバイアスを同時に満足する一つのモデルとして以下のようなものが考えられる。

$$S0(q|p) = \frac{a + d}{a + b + c + d} \quad (7)$$

ただし、 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、あるいは共起確率を表す。 $S0(\bar{q}|p)$ の値は、上式の右辺の a と b, c と d をおのおの入れ替えたものであり、したがって $S0(q|p) + S0(\bar{q}|p) = (a + d)/(a + b + c + d) + (b + c)/(b + a + d + c) = 1$ が成立する。また $S0(p|q)$ の値は、 b と c を入れ替えたものであり、 $S0(p|q) = (a + d)/(a + c + b + d) = S0(q|p)$ が成立する。さらに $S0(\bar{q}|\bar{p})$ の値は、 a と d, b と c を入れ替えたものであり、 $S0(\bar{q}|\bar{p}) = (d + a)/(d + c + b + a) = S0(q|p)$ が成立する。

上述のモデルでは、 $S0(q|p)$ の値にかかわらず常に対称性と相互排他性が成立する。しかし、そもそも p と q の間に何の相関も認められないような場合にも、対称性や相互排他性が成立するとは考えにくい。たとえば P.C. Wason によって見いだされた「Wason 選択課題」[Wason 66] に関する実験結果からも人間が対称性バイアスを持つことが示唆されるが、その一方で完全に対称性が成立する、すなわち人間が条件文を完全に双条件的に捉え「同意 (equivalent)」として解釈すると考えたとしても、被験者実験の結果をうまく説明できない [松井 04]。我々は、否定バイアスを満たすことを前提条件とした上で、対称性と相互排他性の両方を緩く含むモデルが高いパフォーマンスを出せると予想し、それを模索する試行錯誤の結

図 1 S モデルに見られる対称性バイアス

図 2 S モデルに見られる相互排他性バイアス

果, 次のような式が狙いを満たすものであることを見いだした.

$$S(q|p) = \frac{a + bd/(b+d)}{a + bd/(b+d) + b + ac/(a+c)} \quad (8)$$

$$S(\bar{q}|p) = \frac{b + ac/(a+c)}{b + ac/(a+c) + a + bd/(b+d)} \quad (9)$$

$$S(p|q) = \frac{a + cd/(c+d)}{a + cd/(c+d) + c + ab/(a+b)} \quad (10)$$

$$S(\bar{q}|\bar{p}) = \frac{d + ac/(a+c)}{d + ac/(a+c) + c + bd/(b+d)} \quad (11)$$

本稿では, このモデルを対称性モデル (以下 S モデル) と呼ぶ.

図 1 に, a, b, c, d に, $a + b + c + d = 1, 0 \leq a, b, c, d \leq 1$ を満たす様々な乱数データを与えた場合の $S(q|p)$ と $S(p|q)$ の関係を示す. ただしデータは 10000 回分与えた. 図から, $S(q|p)$ の値が大きくなればなるほど, $S(q|p) = S(p|q)$ という関係が成立する傾向性にあることが確認できる. また図 2 に, 図 1 と同じデータを与えた場合の $S(q|p)$ と $S(\bar{q}|\bar{p})$ の関係を示す. 以上に示されたように, S モデルは, S_0 に比べて緩やかなバイアスを持つ.

このモデルが人間の感覚にどれだけ合致しているのかを見るため, [Hattori 03] において行われた被験者実験の

表 1 人間の評価値とモデル出力値の比較

a	b	c	d	Human	H	ΔP	S_0	S
1	1	12	28	11.5	.20	.20	.69	.51
1	0	24	24	11.1	.20	.50	.51	.51
1	0	19	78	11.8	.22	.80	.81	.51
3	1	6	5	39.0	.50	.20	.53	.56
3	1	6	18	31.2	.50	.50	.75	.57
1	0	3	12	29.4	.50	.80	.81	.57
12	3	3	2	64.1	.80	.20	.70	.71
12	3	3	7	69.2	.80	.50	.76	.72
9	1	3	27	75.6	.82	.80	.90	.75

データとの適合性を調べた. この実験では, 原因となる事象 p を「ミルクを飲む」, また結果となる事象 q を「腹痛」とし, $p, \bar{p}$ と $q, \bar{q}$ の間の共起情報を被験者に与える. 被験者は, 与えられた情報をもとに, p と q の因果関係の強度を 100 段階で評価する. 被験者に与えられた共起情報 (a, b, c, d) と被験者 50 人の評価値の平均 (Human) を, [Hattori 03] から引用して表 1 に示す. ここで a, b, c, d はおのの $pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}$ の共起頻度を表す. また表の右側に, これらの共起情報を $H, \Delta P, S_0, S$ に与えた場合の出力結果も併せて示す. ただし H を出力する際には [Hattori 03] にしたがって, $P(q|p) = a/(a+b), P(p|q) = a/(a+c)$ と置き, $H(q|p) = a/\sqrt{(a+b)(a+c)}$ として計算を行った. ΔP についても同様に, $\Delta P(q|p) = (ad - bc)/(a+b)(c+d)$ として計算を行った.

各被験者には, 表に示された 9 種類の実験が与えられた (各行が一つの実験に対応する). ここで a, b, c, d の値は, H および ΔP の出力値が, 低 (0.2), 中 (0.5), 高 (0.8) の組み合わせになるよう恣意的に決められている. 低, 中, 高の組み合わせが 9 通り存在するため, ここでは 9 種類の実験が行われている^{*4}. 各実験において, 被験者は表の各行 (a, b, c, d) に示された頻度で, 対応する共起事象を見せられる. ただし見せられる順番はランダムである. そして実験の最後に p と q の因果関係の強度を評価する. これを 9 種類の実験全てに対して繰り返す.

この被験者実験に関する限り, 人間の評価値と ΔP および S_0 の間に相関は見られない (おのの $r^2 = 0.0, 0.16$). 一方 H と S の出力結果は, 人間の評価値と非常に強い相関を示す (おのの $r^2 = 0.96, 0.97$).

4. N 本腕バンディット問題

上述のように, 人間は対称性バイアスや相互排他性バイアスなど, 非論理的推論を行う傾向性を持つ. 論理的に間違っているにもかかわらず, なぜこのようなバイア

^{*4} この実験は, DH モデルと ΔP モデルの性能を比較するために行われた. H と ΔP の出力値が同じになるような入力 (共起情報) のみを与えても, モデル間の性能比較にならない. このため, このような組み合わせが選ばれている.

スを適用し、事象間の関係を因果関係として捉えるのだろうか。この問題について考察するため、強化学習の分野で「2本腕バンディット問題」と呼ばれる以下のような課題を考える [Sutton 98]。

A と B という二つの餌場があり、主体は毎回どちらかの餌場を訪問するとしよう。その際主体は、餌場毎にある決まった確率で報酬が与えられる（餌場毎に決められた報酬確率をおのおの $\hat{P}(\text{報酬}|A)$, $\hat{P}(\text{報酬}|B)$ と表記する）。ただし、どちらの餌場を訪問するかによって、報酬の得られる確率は異なるものとする。すなわち $\hat{P}(\text{報酬}|A) \neq \hat{P}(\text{報酬}|B)$ である。主体はより多くの報酬を獲得するために、毎回どちらかの餌場を訪問した後、「どちらの餌場を訪問すれば、より多くの報酬を獲得できるか」を判断し、次の訪問先を決定する。この課題における主体の目的は、ある一定の期間内に、なるべく多くの報酬を獲得することである。すなわち主体は、できるだけ速やかかつ正確に、より多くの報酬を獲得できる餌場を推定することが求められる。

この課題に関して以下の二点に注意されたい。第一は、主体は初期段階において餌場に関する情報は何も持たず、したがって自ら情報収集を行わなければならない、という点である。正しい推定を行うためには、両方の餌場についてなるべく多くの情報を収集しなければならない。しかし同時に、主体はより多くの報酬を獲得しなければならない。つまり、餌場の訪問という行為は、情報の獲得と報酬の獲得という二つの目的を担う。一般にこれらの目的は相反する。各餌場についてより多くの情報を収集するためには、両方の餌場をまんべんなく探索する必要がある。一方 $\hat{P}(\text{報酬}|A)$ と $\hat{P}(\text{報酬}|B)$ のどちらが大きいか、すなわち正解は予めどちらかに決まっている。もし主体が情報収集を優先し、両方の餌場をランダムに訪問し続けるならば、正解率は $1/2$ に留まり続ける。しかしだからといって、少ない情報から推定した知識に基づいてどちらか一方を訪問し続けたとしても、その選択は間違っている可能性が高い。この問題は、探索と知識利用のトレードオフと呼ばれる [Sutton 98]。この問題に対して、経験の少ない初期段階には多くの探索をし、知識が蓄えられるにしたがって知識利用の割合を増やしていくという対策が考えられよう。ただしこの場合、探索と知識利用の割合を、時間とともにどのように調整していけばよいのかが問題となる。

第二に注意したいのは、この課題における主体の目的は、 $\hat{P}(\text{報酬}|A)$ や $\hat{P}(\text{報酬}|B)$ の値そのものを推定することではなく、両者のどちらが大きいかを推定することである、という点である。ここで「餌場 A を訪問する」と「餌場 B を訪問する」をおのおの $p, \bar{p}$ と表記することにしよう。また「より多くの報酬を得る」と「より少ない報酬を得る」をおのおの $q, \bar{q}$ と表記する。この時「 p ならば q (A を訪問するならば、より多くの報酬を得る)」がもし真であるならば、そのことから「 $\bar{p}$ ならば $\bar{q}$ (B を

表 2 エージェントによって記憶された共起情報

	餌場 A (p)	餌場 B ($\bar{p}$)
報酬有 (q)	a	c
報酬無 ($\bar{q}$)	b	d

訪問するならば、より少ない報酬を得る)」や「 q ならば p (より多くの報酬を得るならば、A を訪問する)」を導くことは正しい推論である。このように比較に基づいて何らかの判断をしなければならないような状況では、対称性バイアスや相互排他性バイアスは合理的な判断を導く。このことから、これらのバイアスを持つ主体は、持たない主体に比べてこのような課題に対して、うまく対処することが期待される。

次章ではバイアスを実装したエージェントと実装していないエージェントに、上述の課題を与え、シミュレーションによってそのパフォーマンスを調べることにする。

5. シミュレーション

本章では、条件付確率にしたがって推論を行うエージェント (*AgentP* と呼ぶ) と、バイアスを組み込んだ三種類のエージェント (おのおの *AgentS*, *AgentH*, *AgentDP* と呼ぶ) に 2 本腕バンディット問題を与え、そのパフォーマンスを評価する。

5.1 エージェントモデル

各エージェントは、餌場と報酬の有無に関する共起頻度を記憶することができる (表 2)。この表で、例えば a は、これまでにエージェントが餌場 A を訪問し、かつ報酬を獲得できた頻度を表す。また餌場 A, B を訪問することをおのおの $p, \bar{p}$ と、報酬の有無を $q, \bar{q}$ と表記することにする。条件付確率に従う *AgentP* は、これらの情報から、 p と q および $\bar{p}$ と q の関係に関して以下のような信念を形成する。

$$B(q|p) = \frac{a}{a+b} \quad (12)$$

$$B(q|\bar{p}) = \frac{c}{c+d} \quad (13)$$

一方 S モデルに従う *AgentS* は、以下の式を用いて信念を形成する。

$$B(q|p) = \frac{a+bd/(b+d)}{a+bd/(b+d)+b+ac/(a+c)} \quad (14)$$

$$B(q|\bar{p}) = \frac{c+bd/(b+d)}{c+bd/(b+d)+d+ac/(a+c)} \quad (15)$$

同様に DH モデルに従う *AgentH* は以下の式を用いて信念を形成する。

$$B(q|p) = \frac{a}{\sqrt{(a+b)(a+c)}} \quad (16)$$

$$B(q|\bar{p}) = \frac{c}{\sqrt{(a+c)(c+d)}} \quad (17)$$

さらに再定義された ΔP モデルに従う $AgentDP$ は以下の式を用いて信念を形成する．

$$B(q|p) = \frac{1}{2} \left(\frac{ad - bc}{(a+b)(c+d)} + 1 \right) \quad (18)$$

$$B(q|\bar{p}) = \frac{1}{2} \left(\frac{bc - ad}{(a+b)(c+d)} + 1 \right) \quad (19)$$

つぎに各エージェントは、グリーディ方策によって $B(q|p)$ と $B(q|\bar{p})$ の値の大きい方の餌場を選択する [Sutton 98]．ここで各エージェントが選択する（訪問する）餌場を $Action$ と表記することにすれば、

$$Action = \begin{cases} A, & B(q|p) \geq B(q|\bar{p}) \\ B, & otherwise. \end{cases} \quad (20)$$

である．

前章で述べたように、2 本腕バンディット問題におけるエージェントの目的は、できるだけ速やかかつ正確に、より多くの報酬を獲得できる餌場を推定することである．したがって各エージェントの性能を評価する際、例えば正解率のような尺度で表される「正確さ」のみを比較しても正当な評価にはならない．また同じエージェントであっても、探索と知識利用の割合を変えることによって正解率は変わってしまう．

この課題に関しては、「速さ」と「正確さ」を同時に評価する必要があるが、両者の間には、探索と知識利用の割合を変えることによって、一方の性能を良くすれば、他方の性能は悪くなるというトレードオフが存在すると予想される．各エージェントの性能を評価する一つの方法として、あるエージェントに見られるトレードオフを、別のエージェントが破ることができるかをみる、という方法も考えられよう．

このような評価を行うために、上述の 4 体以外にも、探索と知識利用の割合を時間とともに調整するエージェントを導入する．このエージェントは、信念の形成に関しては $AgentP$ と同様であるが、餌場を選択する際に、グリーディ方策ではなく ϵ グリーディ方策を用いる [Sutton 98]．ここで ϵ グリーディ方策とは、 ϵ の確率で餌場 A か B かをランダムに選択し、 $1 - \epsilon$ の確率でグリーディ方策を用いるというものである．このエージェントモデルでは、 ϵ の値を一定とするのではなく、以下の式を用いて時間とともに値を減少させていく．

$$\epsilon = \frac{0.5}{1.0 + \tau t} \quad (21)$$

ここで τ は、 ϵ の値を減少させる度合いを制御するパラメータであり、 τ が大きければ大きいほど ϵ は急速に減少する．また t は、エージェントが選択を行った回数（ステップ数）である．以上のようなエージェントを $AgentP(\tau)$ と呼ぶことにする．ここで $\tau = \infty$ とおけば、 $t = 0$ の場合を除いてグリーディ方策に従うことになり、 $AgentP$ と同様の振る舞いをする．つまり、 $AgentP$ は $AgentP(\tau)$ の特殊な場合といえる．

図 3 正解率の時間発展 ($AgentS$, $AgentH$, $AgentP$, $AgentDP$)

各エージェントは、まず、記憶している共起情報から、各餌場を選択した時に報酬が得られるか否かに関する信念を形成する．つぎに、信念に基づき、グリーディあるいは ϵ グリーディ方策によって、訪問する餌場を決定する．餌場を訪問した際に、ある確率で報酬が与えられ、記憶している共起情報を更新する．このような過程を毎ステップ繰り返す．

5.2 シミュレーション 1

本シミュレーションの目的は、第一に $AgentP$ および $AgentP(\tau)$ を用いて、探索と知識利用のトレードオフを検証すること、第二にこのトレードオフを破ることができるかをみることによって、各エージェントの性能を評価することである．

§1 設定

$AgentS$ と $AgentH$, $AgentP$, $AgentDP$ の 4 体に加えて $AgentP(\tau)$ (ただし τ は 0.05 ~ 1.0 の間で 0.05 刻みで与えた) 20 体の計 24 体に対して、ランダムに生成した 100000 個の 2 本腕バンディット問題を与える．ただし、餌場 A で報酬が与えられる確率 $\hat{P}(q|p)$ は、以下の方法で問題毎に決定した^{*5}．

$$\hat{P}(q|p) = \frac{\sum_{i=1}^n rnd_i}{n} \quad (22)$$

ここで rnd_i は $[0.0, 1.0]$ 区間の一様分布から抽出した乱数である．また $n = 6$ とした．餌場 B において報酬が与えられる確率 $\hat{P}(q|\bar{p})$ に関しても同様に、上式を用いて決定した．

§2 結果と考察

図 3 に $AgentS$, $AgentH$, $AgentP$, $AgentDP$ の正解率の時間発展を示す．ただし各ステップの正解率とは、100000 回の試行のうちで、エージェントが $\hat{P}(q|p)$ と $\hat{P}(q|\bar{p})$ の値の大きい方の餌場を選ぶことができた割合である．なお、シミュレーションは 2000 ステップまで

*5 (22) 式は、 $[0.0, 1.0]$ 区間において、近似的に平均 0.5 の正規分布を生成する．ただし分散は n によって決まる．

図 4 正解率の時間発展 ($AgentP(\tau)$) 図 5 $AgentP(\tau)$ に見られるトレードオフと $AgentS$ によるその解消

行ったが^{*6}、初期段階における振る舞いを詳細に見るため、図では 50 ステップまでを表示している。この図から、 $AgentH$ や $AgentDP$ 、 $AgentP$ に比べて $AgentS$ の性能が圧倒的に優れていることがわかる。

また図 4 に $AgentP(\tau)$ の正解率の時間発展を示す。ただし、ここでは見やすさのため $\tau = 0.05, 0.25, 1.0$ の 3 体の結果のみを表示する。この図から、 τ の値が小さいほど、すなわち探索の割合を大きくすればするほど、正解率の立ち上がりが悪くなる一方で、最終的に達成できる正解率は高い傾向にあることが見て取れる。

この点を詳しく見るため、図 5 に $AgentP(\tau)$ の初期 5 ステップ (1~5 ステップ) における正解率の平均値と最終 5 ステップ (1996~2000 ステップ) における正解率の平均値の関係を示す。また比較のため、 $AgentS$ 、 $AgentH$ 、

$AgentP$ 、 $AgentDP$ のデータも示す。なお、この図には $AgentP(\tau)$ のデータに対する回帰直線 ($r^2 = 0.999$) も併せて表示している。この図から、 $AgentP(\tau)$ および $AgentP$ は、初期段階の正解率が高ければ高いほど、最終的な正解率は低くなる傾向を持つことがわかる。このように、条件付確率にしたがうエージェントには、探索と知識利用のトレードオフが見られる。 $AgentH$ のデータは、回帰直線の左下にある。また $AgentDP$ のデータは回帰直線のわずかに右上にある。一方、 $AgentS$ のデータは、回帰直線のはるか右上にある。すなわち、このエージェントはトレードオフを解消する能力を持つといえる。

以上のように $AgentS$ は 2 本腕バンディット問題に対して、 $AgentH$ や $AgentP$ 、 $AgentDP$ よりも優れた性能を発揮する。以下ではその理由について分析していく。

5.3 シミュレーション 2

シミュレーション 1 において、2 本腕バンディット問題に対する $AgentS$ の優位性が示された。ただし同じ 2 本腕バンディット問題であっても、両方の餌場の報酬確率が近い場合や離れている場合などがあり得る。本シミュレーションの目的は、どのような問題において特に $AgentS$ の優位性が見られるのか、また、なぜ / どのようにして $AgentS$ がトレードオフを破ることができるのか、などについて分析することである。

§ 1 設定

前節のシミュレーションにおいて、各餌場で報酬を与えられる確率 $\hat{P}(q|p)$ 、 $\hat{P}(q|\bar{p})$ は問題毎にランダムに決められた。これらの値が近ければ近いほど、両者のどちらが大きいのかの推定は困難になる。ここで問題の「難易度」を表す尺度として、 $1 - |\hat{P}(q|p) - \hat{P}(q|\bar{p})|$ を考える。この値は $[0.0, 1.0]$ 区間の実数となるが、この区間を 10 等分し、問題の難易度を 10 段階で表す。シミュレーション 2 では、各難易度の問題を 10000 個ずつ作成し、計 100000 個の問題を $AgentS$ 、 $AgentH$ 、 $AgentP$ 、 $AgentDP$ に与える。なお初期段階におけるエージェントの振る舞いを分析するため、このシミュレーションは 200 ステップまで行う。

§ 2 結果と考察

図 6 に問題の難易度と各問題に対するエージェントの探索率の関係を示す。ただし探索率とは、シミュレーションの期間内に、エージェントが訪問先を変更した割合のことである。図で、横軸は難易度を表す。縦軸は、横軸で表される難易度の問題が与えられた場合、各エージェントがどの程度の割合で探索を行ったのかを表す。ただし、ここでは難易度毎に 10000 個の問題に対する探索率の平均値を示している。この図から、 $AgentS$ は問題が難しくなればなるほど、情報収集のためにより多くの探索を行うようになることがわかる。特に、難易度が 6 以上の問題に対して、その傾向が顕著である。一方 $AgentH$ や $AgentP$ 、 $AgentDP$ には、このような傾向はわずかに

^{*6} 各エージェントの正解率は、2000 ステップまでにほぼ収束する。1900 - 2000 ステップ間における正解率の平均上昇率 (t ステップ目における正解率を $CorrectRate(t)$ と表記するとすれば、 $(CorrectRate(2000) - CorrectRate(1900)) / 100$) は、全てのエージェントにおいて $1.0E-5$ 未満となる。ちなみに、上昇率の最も高いエージェントは $AgentS$ であった。したがって、シミュレーション期間を延長したとしても、 $AgentS$ の優位性は保たれると考えられる。以上の理由により、シミュレーションは 2000 ステップまでで打ち切った。

図 6 難易度に応じた探索率の調整

 図 8 $AgentS$ の内部状態 (難易度 1)

図 7 難易度に応じた正解率

 図 9 $AgentS$ の内部状態 (難易度 10)

認められる程度である。

図 7 に、問題の難易度と正解率の関係を示す。ただし縦軸の正解率は、シミュレーション期間を通しての正解率の平均値を示す。この図から、どのような難易度の問題に対しても、 $AgentS$ の正解率が最も高いことが確認できる。もちろんその正解率は、他のエージェントの場合と同様、問題が困難になるにつれて下降する。この図と図 6 を比べてみると、 $AgentS$ の探索率は、正解率の下降に伴って急激に上昇することがわかる。さらに、図からはわかりにくいだが、難易度が 3 以下の簡単な問題に関しては、 $AgentS$ の探索率が $AgentP$ や $AgentDP$ のそれを下回るにもかかわらず、より高い正解率を達成する。これは $AgentS$ が、簡単な問題に対してはより少ないサンプリングによって、より正しい推論を行えることを意味する。

以上の結果から、問題の難易度に応じて、 $AgentS$ の探索と知識利用の比率が自律的に調整されることによって、高い正解率を達成していることがわかる。この調整能力が、探索と知識利用のトレードオフを破る要因となっていると考えられる。

図 8 と図 9 に $AgentS$ の内部状態を示す。図 8 は、難易度 1 の問題を与えた場合の最終ステップ (200 ステッ

プ目) における $B(q|p)$ と $B(\bar{q}|\bar{p})$ の関係を示している。なおシミュレーションは 10000 試行行っているので、図には 10000 個のデータがプロットされている。 $B(q|p)$ と $B(\bar{q}|\bar{p})$ の間には、(5) 式で示したような正確な比例関係は成立しないが、 $B(q|p)$ が大きくなれば $B(\bar{q}|\bar{p})$ も大きくなるという相互排他性バイアスが見て取れる。なお図 8 と図 9 は、 $B(\bar{q}|\bar{p}) = 1 - B(q|p)$ に関してほぼ対称な形をしている。これは、本稿では便宜上、餌場 A を p 、餌場 B を $\bar{p}$ とおいたが、両者を逆転させても結果は変わらないということを意味する。以後、 $B(\bar{q}|\bar{p}) \geq 1 - B(q|p)$ の範囲にあるデータを対象に分析を行う。

図 9 には、難易度 10 の問題を与えた場合の内部状態を示す。 $B(q|p)$ の値が 1.0 付近にある時には、難易度 1 の問題の場合と同様、 $B(q|p)$ が大きくなれば $B(\bar{q}|\bar{p})$ も大きくなるという傾向性が見られる。しかし、それ以外の領域、特に $B(q|p) < 0.5$ では、状況が一変する。興味深いことに、この領域においては、 $B(q|p)$ の値が小さくなればなるほど $B(\bar{q}|\bar{p})$ の値が大きくなるという、いわば逆相互排他性バイアスが見られる。 $AgentS$ ではモデルの設定上、否定 $B(q|\bar{p}) = 1 - B(\bar{q}|\bar{p})$ が成立する。したがって逆相互排他性バイアスが見られるということは、 $B(q|p)$ と $B(q|\bar{p})$ が一致する傾向にあるということを示

図 10 最大報酬確率と $|B(q|p) - B(q|\bar{p})|$ の関係 (難易度 10)

味する。

逆バイアスに関してさらに分析を進めよう。まず難易度 10 の問題とは、各餌場で報酬が与えられる確率 $\hat{P}(q|p)$ と $\hat{P}(q|\bar{p})$ が近い問題であった。ただし、これらの値それぞれは様々な値を取る。ここで、二つの餌場の最大報酬確率 $\max(\hat{P}(q|p), \hat{P}(q|\bar{p}))$ を考える。図 10 に、難易度 10 の 10000 個の問題に関して、最大報酬確率と $|B(q|p) - B(q|\bar{p})|$ の関係を示す。図から、最大報酬確率が 0.5 以下の領域で、 $B(q|p)$ と $B(q|\bar{p})$ が一致する傾向にあることがわかる。つまり、図 9 において逆バイアスが観察されたのは、難易度 10 の問題のうち、最大報酬確率が 0.5 以下の問題に対するデータであったことがわかる。逆相互排他性バイアスが生じる問題領域では、どちらの餌場が正解なのかを推定することが困難である上、どちらを選択しても低い確率でしか報酬を得られない。

さて難易度が 5 以下の問題とは、 $\hat{P}(q|p)$ と $\hat{P}(q|\bar{p})$ の差が 0.5 以上あるような問題である。すなわち、これらの問題においては、一方の餌場の報酬確率が 0.5 以上ならば、他方は 0.5 以下であることが保障されている。このような保障がある場合には、相互排他性バイアスは合理的な推論を導く。しかし、いつもこのような関係が満たされるわけではない。難易度が 5 以上の問題、例えば難易度 10 の問題では、 $\hat{P}(q|p)$ と $\hat{P}(q|\bar{p})$ の値が非常に近い。すなわち、難易度の高い問題では、「 $\hat{P}(q|p)$ が大きい (小さい) ならば $\hat{P}(q|\bar{p})$ も大きい (小さい)」という関係が成立する。これは簡単な問題とは逆の関係である。このような問題に対して、むやみに相互排他性バイアスを用いることは、誤った推論を導く。すなわち相互排他性バイアスの適用は、問題に応じて適切に調整されなければならない。

ここで各エージェントの相互排他性バイアスの適用具合を見るため、バイアスの強度を表す尺度として相関係数を用いることにする。相互排他性バイアスとは、一方の餌場の報酬確率が高いと信じる時、他方の報酬確率は低いと信じる傾向性である。このため、報酬確率が高いと信じている餌場を x とする時、 $B(q|x)$ と $B(q|\bar{x})$ の相

図 11 難易度に応じた相互排他性バイアスの強度

関係数を求め、これを相互排他性バイアスの強度とする。

図 11 に、難易度別の相互排他性バイアスの強度を示す。AgentDP はモデルの設定上、相互排他性を持つため常に 1 の値を取る。一方 AgentP はモデルの設定上は、相互排他性バイアスも逆バイアスも持たない。しかし難易度の高い問題では、「 $\hat{P}(q|p)$ が大きい (小さい) ならば $\hat{P}(q|\bar{p})$ も大きい (小さい)」という関係が成立するため、それを反映して難易度が高くなるにしたがって逆バイアスを持つようになる。AgentS は簡単な問題に対しては相互排他性バイアスを適用するが、その強度は、問題が難しくなるにつれて小さくなる。特に難易度 7 以上の問題に対しては、その値が負となる。これは、AgentS は難易度の低い問題に対しては相互排他性バイアスを適用するが、難易度の高い問題に対しては、逆バイアスを適用すること、つまり難易度に応じてバイアスの適用を調整していることを示している。

以上に示したように、AgentS は緩やかな相互排他性バイアスを持つ。ここで「緩やか」とは、このバイアスが有効に作用するような難易度の低い問題に対してはバイアスが適用されるが、そうでない場合はむしろ逆バイアスが適用される、という意味である。AgentS においてはその緩やかさゆえ、簡単な問題が与えられた場合には、 $B(q|p)$ と $B(q|\bar{p})$ の値の差がより増幅される一方、難易度の高い問題が与えられた場合には、 $B(q|p)$ と $B(q|\bar{p})$ の値がより接近する。両者の値が近づけば近づくほど、(20) 式に基づいて行われるエージェントの行動選択は、少しのゆらぎによって変動しやすくなる。これが、図 6 で示した探索率調整の要因になっていると考えられる。

6. まとめと今後の課題

本稿では、人間に見られる非論理的推論の事例として、対称性バイアスと相互排他性バイアスを取り上げた。そして、これらのバイアスを因果推論の観点からモデル化した。また、非論理的推論の有効性を確かめるため、モデル化した推論機構をエージェントに組み込み、2 本腕

バンディット問題においてその性能を評価した。

1 章で述べたように、対称性や相互排他性バイアスは「 p ならば q 」から「 q ならば p 」や「 $\bar{p}$ ならば $\bar{q}$ 」を推論する傾向性と捉えられる。つまり、これらのバイアスを持つ主体は「 p ならば q 」を信じる時「 q ならば p 」や「 $\bar{p}$ ならば $\bar{q}$ 」の真偽が定かではないにもかかわらず、それらを信じてしまう、いわば“せっかち”な主体であるとみなすことができる。一方、この主体を「 p ならば q 」を因果関係として捉えてしまう主体とみなせば、 $P(q|p)$ が高くても、 $P(p|q)$ や $P(\bar{q}|\bar{p})$ が低ければ、なかなか「 p ならば q 」とは信じない、“慎重”な主体であるとみなすこともできる。このように、バイアスを持つ主体は、“せっかち”と“慎重”という二面性を持つことになる。

このような二面性は論理的に誤った推論を導くため、 p と q の共起情報から条件付確率 $P(q|p)$ を正確に推定しなければならないような課題、例えば天気予測のような課題に対しては、マイナスの効果しか持たない。しかし N 本腕バンディット問題のように、比較に基づいて何らかの判断を行うような課題に関しては、合理的な推論を導く。ただしこの課題には、探索と知識利用のトレードオフと呼ばれる問題がある。5・3・2 節で見たように、本稿で提案したエージェントは、より多くの情報を必要とする難しい課題に対しては、探索比率が上昇することで高い正解率を達成する。これは、バイアスの“慎重”な側面が強調された結果といえる。一方簡単な課題に関しては、知識利用を優先することで、少ない情報から速やかに正解を導く。これは、“せっかち”な側面が強調された結果である。このように、提案エージェントでは、“慎重”と“せっかち”の二面性が必要に応じて調整されることで、効率的な問題解決がはかれる。さらに提案エージェントに、比較が困難であり、かつどちらを選択しても損をするような課題を与えた場合には、逆相互排他性バイアスという興味深い現象も観察された。

本稿では 2 本腕バンディット問題において、提案エージェントの性能を評価したが、以上に示した能力は、言語学習のような課題に対しても有効であると思われる。1 章で述べたように、相互排他性バイアスとは、対象は一つのラベルのみを持つと考える傾向性のことである。このバイアスは、同義語を回避しようとする傾向性と解釈することもできる [Smith 04]。しかし、我々の言語には同義語が数多く存在する。また同義語でなくとも、「犬」と「ポチ」のように二つのラベルが同一の対象を指示する場合もあり得る。相互排他性バイアスは、このようなラベルの学習を阻害する [Markman 90, 針生 00, 今井 01]。つまりバイアスを有効に作用させるためには、むやみに適用するのではなく、状況に応じてその適用を抑制する必要がある。

今、二つのラベル l_1, l_2 と二つの対象 o_1, o_2 があってしよう。ここでエージェントが、各ラベルの指示対象を同定しようとしている状況を想定する。もし l_1 と l_2 が

同義語ではなく、おのれの o_1 と o_2 を指示するとすれば、相互排他性（同義語回避）バイアスを適用することが望ましい。例えば「 l_1 は o_1 である」という知識がある時、そこから「 l_2 は o_2 である」を推論することは正しい結果を導く。一方 l_1 と l_2 が同義語であり、ともに o_1 を指示するとすれば、バイアスの適用は誤った結果を導く。もちろんエージェントは、前もって二つのラベルが同義語か否かの情報を持っているわけではない。エージェントは、このような状況で各ラベルの指示対象を推定しなければならない。

二つのラベルが同義語でない場合、学習過程において o_1 を見た時に l_1 を聞く確率は、 l_2 を聞く確率よりも高くなるだろう。つまりこの状況で、対象とラベルの対応関係を推定することは、前章で述べた難易度の低い問題になる。一方二つのラベルが同義語である場合、両者の確率値は近くなり、難易度の高い問題になる。したがって、提案エージェントは、二つのラベルが同義語でない場合には相互排他性バイアスを適用し効率的な言語学習を行う一方、そうでない場合は、逆相互排他性バイアスを適用することで、二つのラベルを同義語とみなすことが期待できる。このような逆相互排他性バイアスは幼児においても観察されている。相互排他性バイアスは、生後 17 ヶ月程度で見られるようになる。ただし、それ以前にはこのバイアスを単に持たないのではなく、生後 14 ヶ月の時期には、新奇なラベルを既知の対象の名前と解釈するという逆バイアスを持つとの報告がある [Halberda 03]。この報告との整合性についての分析、および提案モデルの言語学習課題への適用は今後の課題としたい。

謝 辞

本研究の一部は、文部科学省 21 世紀 COE プログラム「インテリジェントヒューマンセンシング」の補助による。

◇ 参 考 文 献 ◇

- [Davidson 05] Davidson, D. and Tell, D.: 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 (2005)
- [藤田 01] 藤田 和生：霊長類のコミュニケーション, 乾 敏郎, 安西 祐一郎 (編), コミュニケーションと思考, pp. 79–114, 岩波書店, 東京 (2001)
- [Goswami 98] Goswami, U.: *Cognition in Children*, Psychology Press, London (1998)
- [Halberda 03] Halberda, J.: The development of a word-learning strategy, *Cognition*, Vol. 87, pp. B23–B34 (2003)
- [針生 00] 針生 悦子, 今井 むつみ：語意学習メカニズムにおける制約の役割とその生得性, 今井 むつみ (編), 心の生得性 - 言語・概念獲得に生得的制約は必要か -, pp. 131–171, 共立出版, 東京 (2000)
- [服部 01] 服部 雅史：因果帰納の二要因ヒューリスティクス・モデル, 認知科学, Vol. 8, pp. 444–453 (2001)
- [服部 02a] 服部 雅史：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 (2002)
- [服部 02b] 服部 雅史：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 (2002)

- [Hattori 03] Hattori, M.: 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 (2003)
- [Imai 94] Imai, M., Gentner, D., and Uchida, N.: Children's theory of word meanings: The role of shape similarity in early acquisition, *Cognitive Development*, Vol. 9, pp. 45–75 (1994)
- [今井 01] 今井 むつみ: 概念発達と言語発達における類似性の役割, 大西 仁, 鈴木 宏昭 (編), 類似から見た心, pp. 148–178, 共立出版, 東京 (2001)
- [Jenkins 65] Jenkins, H. M. and Ward, W. C.: Judgment of contingency between responses and outcomes, *Psychological Monographs*, Vol. 79, pp. 1–17 (1965)
- [Markman 88] Markman, E. M. and Wachtel, G. F.: Children's use of mutual exclusivity to constrain the meaning of words, *Cognitive Psychology*, Vol. 20, pp. 121–157 (1988)
- [Markman 89] Markman, E. M.: *Categorization and naming in children: Problems of induction*, MIT Press, Cambridge (1989)
- [Markman 90] Markman, E. M.: Constraints children place on word meanings, *Cognitive Science*, Vol. 14, pp. 121–157 (1990)
- [Markman 03] Markman, E. M., Wasow, J. L., and Hansen, M. B.: Use of the mutual exclusivity assumption by young word learners, *Cognitive Psychology*, Vol. 47, pp. 241–275 (2003)
- [松井 04] 松井 理直: 計算論的関連性理論に基づく反事実条件文の解釈, *Theoretical and Applied Linguistics at Kobe Shoin*, Vol. 7, pp. 83–101 (2004)
- [Sidman 82] Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., and Carrigan, P.: A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children, *Journal of the Experimental Analysis of Behavior*, Vol. 37, pp. 23–44 (1982)
- [Smith 04] Smith, K.: The evolution of vocabulary, *Journal of Theoretical Biology*, Vol. 228, pp. 127–142 (2004)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning –An Introduction–*, MIT Press, Cambridge, MA (1998)
- [田中 00] 田中 正之, 松沢 哲郎: シンボルの成立, 渡辺 茂 (編), 心の比較認知科学, pp. 226–268, ミネルヴァ書房, 京都 (2000)
- [Wason 66] Wason, P. C.: Reasoning, in Foss, B. M. ed., *New horizons in psychology*, pp. 135–151, Penguin Books, Harmondsworth, Middlesex, UK (1966)
- [Yamazaki 04] Yamazaki, Y.: Logical and illogical behavior in animals, *Japanese Psychological Research*, Vol. 46, pp. 195–206 (2004)
- [山崎 99] 山崎 由美子: 動物における刺激等価性, 動物心理学研究, Vol. 49, pp. 107–137 (1999)
- [山崎 06a] 山崎 由美子, 日原 さやか, 藤井 直敬, 岡ノ谷 一夫, 入来 篤史: 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 (2006)
- [山崎 06b] 山崎 由美子, 岡ノ谷 一夫, 入来 篤史: 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 (2006)

〔担当委員: 兼岩 憲〕

2006 年 7 月 19 日 受理

著者紹介

篠原 修二

1999 年神戸大学大学院自然科学研究科博士後期課程修了。博士 (理学)。ベンチャー企業で, Web アプリケーションやデータベースシステムの開発に携わった後, 2005 年から豊橋技術科学大学で生体情報を用いた教育支援システムの研究開発に従事。日本認知科学会, 進化経済学会各会員。

田口 亮 (正会員)

2004 年豊橋技術科学大学大学院工学研究科修士課程修了。現在, 同大学院博士後期課程在学中。情報処理学会会員。人間-エージェント対話に興味を持つ。

桂田 浩一 (正会員)

1995 年大阪大学基礎工学部情報工学科卒業。2000 年同大学院基礎工学研究科博士後期課程修了。同年豊橋技術科学大学工学研究科助手。博士 (工学)。マルチモーダル対話, 知識処理に関する研究に従事。AAAI, 情報処理学会, 日本音響学会, 言語処理学会, ヒューマンインタフェース学会各会員。

新田 恒雄 (正会員)

1969 年東北大学工学部電気工学科卒業 (株) 東芝総合研究所, 同社マルチメディア技術研究所を経て, 1998 年豊橋技術科学大学大学院工学研究科教授。工学博士。音声認識・合成・文字認識, マルチモーダル対話システム, および概念獲得の研究に従事。IEEE, 電気情報通信学会, 情報処理学会, 日本音響学会各会員。