

論理と確率についての推理を導く心の中のゲーム

Games in Mind Which Yield Reasoning on Logic and Probability

犬童 健良[†]
Kenryo Indo

[†] 関東学園大学
Kanto Gakuen University
kindo@kanto-gakuen.ac.jp

Abstract

This study presents an application of game theory to the cognitive modeling of human reasoning especially for the four card selection task and for the problem of three prisoners.

Keywords — game theory, four card selection task, problem of three prisoners

1. はじめに

本論文では、ゲーム理論を、認知モデルの設計に応用する。ゲーム理論[18] [11] [13]は複数のエージェントたちの挙動を合理的な選択の観点から説明するための数理モデルである。ゲームの均衡点におけるエージェントのふるまいは、その定義上、人間のような複雑な処理を行うものではなく、他のエージェントの活動に対する予測可能な反応ルールの集まりとして表される。本論文では、エージェントの反応ルールの組は、情報反応モデルと呼ばれ、ゲームモデル設計の仕様となる。少々意外に思われるかもしれないが、ゲームモデルを用いると、必ずしも合理的でない現実の人々の思考を明快に説明することができる。

ゲーム理論は、少なくとも以下に述べる三つの観点において、認知活動を記述する目的に適する。第一に、ゲームとは、一定のルールの下で人々の意識的な目標追求が促されるような状況のことであり、当面の目標と制約（ゲームのルール）が与えられて、各エージェントはそのルールの下で、精一杯、知能をはたらかせると想定される。この意味でのゲームは、実用的な思考のためのスキーマ（ないしフレーム）の総称と考えられる。第二に、ゲームの均衡点はこの意識的な目標追求の到達点として解釈される。ゲーム理論はゲームの均

衡点を見出すマルチエージェントによる並列計算の抽象化である。実際、均衡点を見出す推論を視覚化するラベリングシステム[15]は、知識表現の計算モデルであるフレーム[9]の代用とみなせる。またそれは心の社会[10]に示唆されていたが具体的に示されていなかった定式化を埋める試みでもある。三つ目の観点は、均衡の認知的な解釈である。均衡点是最適反応の組として定義される。その一つの解釈は、人々が環境への適応状態に満足し、それ以上追加的な思考努力をしようとしないう満足された状態[17]である。このように、認知モデルに転用されたゲーム理論には、フレーム理論、心の社会、および満足化原理が弱く結合する。

本論文のアプローチが目指すのは、現実の人間の思考を計算的に近似するゲームを設計する体系立った手法としての、メカニズムデザイン[5]の認知科学版である。またそれは人間の脳による自然の営みをゲームとして人工的に再現する試みでもある。その作法としては、認知的モデラーは、情報反応モデル、ペイオフ表、ラベリング、シミュレーションなどを活用しながら思考活動を表現するゲームを設計する。具体的な応用として、本論文では4枚カードの選択課題[19][20]と3囚人問題[2][3][6]をとりあげる。これらは条件文の論理的意味や確率の計算を問うクイズ問題であり、それぞれ人間の直観的な思考と数学的な正解を導く思考とが鋭く対立することでよく知られている。

ちなみに、ゲーム理論の主な用途は経済的取引や生物・社会の行動分析である。社会的な状況の認知において、ゲーム理論は現実とモデルの自明な対応をもつ。たとえば限られた資源を複数人で争うような状況で無駄のない配分状態（つまりパ

レート最適性)を,ゼロ和ゲームに対応させることは自然である.一方,現実の行為者は必ずしも合理的にふるまわないという別の問題がある.行動経済学や進化ゲーム理論はこの問題を扱うが,本論文の目的から逸れるのでこれ以上言及しない.

本論文の以降の部分では,まず第2節で,ゲームとして認知モデルを構成する方法を,嘘つき文を例に説明する.第3節では,線画の立体解釈の曖昧性を,均衡の安定性と関連させて論じる.第4節では,4枚カードの選択課題における情報反応モデルを説明し,フレーム理論との関係に言及する.否定表現を導入することによる問題のむずかしさの変化は,ペイオフの微妙な変化に応じてデフォルトの均衡が移動することによって説明できる.第5節では3囚人問題をとりあげ,Grice的なコミュニケーションおよびメカニズムデザインの観点で再解釈する.また類題として3ドア問題(モンティホール問題)と夕食問題に触れる.第7節ではラベリングシステムを導入し,3囚人問題を例に,シミュレーション実験を併用してゲームを設計する方法を解説する.ラベリングシステムはゲーム理論によるフレームシステムの別表現とみなされる.最後に第7節でまとめとする.

2. ゲーム理論による認知モデル

本論文では,思考を構成するある要素が特定のシンボルに対応する意識表象として区分できる場合に,その意識的な活動の単位をエージェントと呼ぶ.非協力ゲーム(以下たんにゲームと呼ぶ)では,ゲームをプレイすること以外にコミュニケーションする手段はない.均衡点は各エージェントの最適反応の組である.本論文では,認知活動はゲームとその均衡点によって表現される.

一例として,「文Bは正しくない。」を文A,「文Aは正しくない。」を文Bとする.この意見対立の状況は,表1に示されるような各文のエージェントが「正(T)」か「誤(F)」を選ぶ2人ゲームとして表現できる.

ゲーム1のルール(結果関数)は,文エージェントの選択の組に対するゲームの結果を特定する.表1では行の選択(文Aの真理値)と列の選択(文Bの真理値)の交わるセルに,ゲームの結果として,Aの利得,Bの利得のようにカンマで区切った両者の利得を表示する.

エージェントの戦略には純戦略とランダム戦略(混合戦略)がある.純戦略は確率を用いない戦略であり,ランダム戦略は純戦略を確率で混合する.ゲームの均衡点は,エージェントの選択する戦略(ランダム戦略を含む)が互いに最適反応になっている状態として定義される.均衡点では,各エージェントは他のエージェントの活動を一定とすると自分だけ選択を変える動機をもたない.

ゲーム1における各エージェントの最適反応は以下のような4ルールとして表せる(ただし $\alpha \geq 0$ とし,ランダム戦略を除いて考える).

- ① 文Aが「正」 \Rightarrow 文Bは「誤」,
- ② 文Aが「誤」 \Rightarrow 文Bは「正」,
- ③ 文Bが「正」 \Rightarrow 文Aは「誤」,
- ④ 文Bが「誤」 \Rightarrow 文Aは「正」.

このゲームには,文Aが正しく文Bが誤り(①と④の組),文Bが正しく文Aが誤り(②と③の組)という二つの均衡点があり,それぞれ三段論法によって局所的なループを成す.これらの他にゲーム1にはパラメータ α に依存して両者が確率 p で「正」を選ぶランダム戦略の均衡点がある.たとえば $\alpha = 1/2, 1, 2$ のとき $p = 1/3, 1/2, 2/3$ である.

表1 ゲーム1のペイオフ表

		B	
		T	F
A	T	0, 0	1, α
	F	$\alpha, 1$	0, 0

表2 ゲーム2のペイオフ表

		B	
		T	F
A	T	1, 0	0, 1
	F	0, 1	1, 0

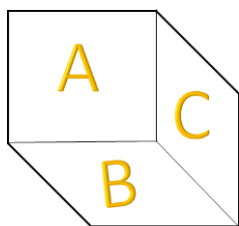


図1 線面の図形

ランダム戦略が許されるとき、ペイオフ表で書けるゲームには均衡点が必ず存在する[11]。表2は、A:「文Bは誤り」とB:「文Aは正しい」の連言に対応するゲームを表している。これは、所謂、嘘つき文であり、純戦略では最適反応全体がループして均衡しない。確率1/2でTを選ぶのが唯一の均衡点である。

3. 認知の曖昧性

もう一つの例として、線面図形の立体知覚（ネッカーの立方体の変種）をモデル化する。図1には図の中心近くで三つの面の交わる点がある。立体図形として見たとき、少なくとも異なる二つの立体知覚が可能である。図を見ている人にとって中央に見える頂点が、周囲の6頂点よりも手前にあるように見える（凸である）か、奥にあるように見える（凹である）かのいずれかの知覚が生じ、両方同時には知覚されない。そこで3面A, B, Cをそれぞれエージェントとみなし、表3にゲーム3としてそのゲームのペイオフ表を与える。

表3の行はAの選択、また列はBとCの選択を結合して示す。各セルはゲーム3の一つの結果に対応し、A, B, Cそれぞれの利得を表す。各行各列の見出し部分のTとFは、エージェントが選択できる各面の中央の点に対する傾きとして読み直して頂きたい。Tは中央の頂点が手前、つまり図を見ている人に近づく方向にその面が傾いていることを表し、Fは中央の頂点が奥、つまり図を見ている人から離れる方向にその面が傾いていることを表す。

表3 ゲーム3のペイオフ表

		B, C			
		TT	TF	FT	FF
A	T	1,1,1	1,1,0	1,0,1	0,1,1
	F	0,1,1	1,0,1	1,1,0	1,1,1

ゲーム3が多数決を表していることは明らかだろう。各エージェントは他の2エージェントが合意しているとき、自分だけ別の解釈をとるのは不利である。ゲーム3には、TTTとFFFの二つの純戦略の均衡点と三者がいずれも確率1/2でTまたはFを選ぶランダム戦略の均衡点がある。前者は凸と凹の立体知覚、後者は平面図としての解釈に対応する。ちなみにネッカーの立方体は隠れている面が交わる辺が図1に追加されたものである（隠れた3面も同じゲームとすると、両ゲームで勝利した向きをゲーム2のプレイとする2段階ゲームとなるが詳しくは読者に委ねる）。

ところで、これまでに見たゲームはいずれも暗黙の了解として同時手番を仮定した。同時手番とは全エージェントが相手の行動を事前に知ることなく自分の行動を選ぶことである。同時手番ではないゲームは展開形ゲームとも呼ばれる。手番を導入することによって、同じペイオフ表の異なる手続き的な意味を考えることができる。たとえばもしジャンケンが同時手番でなかったとしたら、とうぜん後手必勝になる。一方、ゲーム3に手番を導入すると、均衡点は同時手番のときと同じだが、最初の手番のエージェントの独裁となる。

ゲーム3の手番の順序は6通り（同時手番も含むと13通り）あるが、手番の順序によらずTTTかFFFのいずれかが均衡点である。最初の手番のエージェントにとってTとFは無差別である。2番目以降の手番のエージェントの最適反応は、前の手番のエージェントの選択に合意することである。

こうして、図1のような線面を見る人が、3面のうちいずれの一つに注目したとしても、互いに排他的である凸または凹の立体知覚が生じることをゲーム3は予測する。しかしこれらの均衡点は

より強い安定性の条件（部分ゲーム完全性）を満たしていない。実際、仮に2番目のエージェントが1番目と異なる選択をしたとすると、最後の手番でTとFは無差別になる。このように予め均衡点で計画された行動と、実際の選択が異なることは動的矛盾と呼ばれる。一方、線画を見ている人の視点が自由に動くような状況は、ゲーム2における手番を任意にして反復する場合に相当する。視点の移動は、ネッカーの立方体における解釈の曖昧性の要因と考えられるので、均衡点の不安定性はこうした仮説とも矛盾しない。

4. 4枚カードの選択課題

Wasonのカード選択課題[19]では、机の上に並べられた4枚のカードのうち、「もしカードの一方の面が母音なら、別の面には偶数の数字が書かれている」というルールが正しいかどうか確かめる必要のあるカードを選ばせる問題である。この問題でルールを実質的含意 ($p \rightarrow q = \neg p \vee q$) として解釈すると、正しく答えられる被験者の割合は低い（1～2割程度といわれる）。また結論部を否定表現に変えたルール ($p \rightarrow \neg q$) では、正答率は8割程度まで上がることも知られる[1]。

選択問題は情報の下での合理的選択問題として表現することができる[7][12]。とうぜんながら合理的選択の直接の適用は数学的な正解に一致する。事実、報酬と努力の差としての利得最大化（スペルベルらの適切性原理に相当する）は実証的な選択パターンを説明できない。また[12]の情報量最大化でも基本的な等確率の場合を説明できず、稀有仮説[20]のような仮定が置かれる。

本節では、選択課題のパフォーマンスを説明するゲームモデルを構築する。ここでは、2人のエージェントAとCをそれぞれルール $p \rightarrow q$ の条件部 p と結論部 q の真理値に対応させたゲームを考える。ゲームの結果は以下の4通りである。

- ① (T, T) \Rightarrow 違反しない
- ② (T, F) \Rightarrow 違反する
- ③ (F, T) \Rightarrow 違反しない

- ④ (F, F) \Rightarrow 違反しない

以降の部分では、ルール①～④の左辺に現れるゲームの結果を、括弧とカンマを省いて短くし、TT, TF, FT, TTと書くことにする。これらのゲームの結果に応じて、第3のエージェントとしての被験者Sが、最終的に、選択課題に回答すると考える。

ゲームが与えられると、隠れている方（＝ウラ）の選択は見えている方（＝オモテ）に対する最適反応として予測される。もしウラの最適反応に違反のケースが含まれていればそのカードを調べ、さもなければ調べないのがSの最適反応である。

表4として、正答パターンを選択できるようなAとCのゲームの一例を示す。

表4 ゲーム4のペイオフ表

		C	
		T	F
A	T	1, 1	0, 0
	F	1, 1	1, 1

読者は表4に示したゲーム4のペイオフ表が、実質含意の真理表と似ていることに気付かれたであろう。ゲーム4におけるAの最適反応は $T \Rightarrow T$, $F \Rightarrow T \vee F$, Cの最適反応は $T \Rightarrow T \vee F$, $F \Rightarrow F$ であり、均衡点はTT, FF, FTの三つである。選択課題の正答は（T, F）に該当するカードだけを選択する。

均衡FTをデフォルトとすると、ゲーム4は選択課題の論理的な正解（ p と $\neg q$ の選択）を再現する。フレーム理論[9]にしたがって、このゲームを認知モデルとして解釈しよう。まずフレーム理論について簡単に説明する。一つのフレームは質問（終端）の集まりであり、質問の答えを他のフレームと交換し合い、互いに結びつくことで構造体（フレームシステム）を形成する。未回答の質問に弱く結びついた暫定的な答えをデフォルトという。認知過程はフレームの再計算であり、デフォルトはその処理を効率化する。

このように、デフォルトは暫定的な答えを保

持することで、新しい質問の答えが見つかるまでの間、品質の劣る情報を検閲する役割を果たすと考えられる。また興味深いことに、Minsky はデフォルトが文化や言語や制度によって変わることを指摘している。

ここで情報反応モデルについて、直観的な解釈を与えておこう。ルール $p \rightarrow q$ は、実質含意と矛盾せず、「 p の成立を伝えるメッセージを受信したら、ある成果 q を信じてよい」という実用的な意味に解釈しうるが、これは情報反応モデルを、最適反応を用いずに概念的に述べたものである。

デフォルトの均衡 FF は、この解釈の下で、“No Message” かつ “No Fruits” の状態に相当する。また本論文ではデフォルトと矛盾しないメッセージを、(記号論から借用して) 無徴 (unmarked) であると言う。またデフォルトと矛盾し、フレームの再計算を促すメッセージは有徴 (marked) であると言う。要するに、選択課題の情報反応モデルは、有徴である (= 被験者が気になる) カードは調べるが、無徴である (= 被験者が気にならない) カードは調べないと予測する。

さて選択課題に対する現実の人々の選択パターンを再現するゲームを表5に示す(表5参照)。

表5 ゲーム5のペイオフ表

		C	
		T	F
A	T	1, 1	0, 0
	F	1/2, 1	1, 1

ゲーム5では FT のときのAの利得がゲーム4から半減する(あるいは1より少しだけ下げる)。これによって、 FT が均衡点でなくなり、代わりにCの側のTの確率を $2/3$ まで下げたランダム戦略の均衡が一つ生じる。同様に、 FT のCの利得を下げることによって FT が均衡点でなくなるが、両方同時に下げるとゲーム1の最適反応にさらに近づく。もしAの利得が TT でも減っても、 FT よりわずかに上回るか、Cの利得が FF で減っても FT よりわずかに上回るならば、ゲーム5

の情報反応モデルは質的に保たれる。

このようにゲーム4のペイオフを若干変更しただけで、デフォルトの均衡は移動し、 TT か FF のいずれかになる。もし TT がデフォルトなら、 p と q の観察は再計算を生じないので、実証的な選択パターンを再現できない。もし FF がデフォルトであると考えたとすれば、 p と q の観察では再計算され、 $\neg p$ と $\neg q$ では再計算されないから、実証的な選択パターンを再現できる。すなわち、 $\neg p$ や $\neg q$ がカードのオモテ面として観察されたとき、デフォルトである FF に対応する命題 $\neg p \cdot \neg q$ と両立しない。このため均衡点の再計算が起こる。またこのとき戦略組 FF の一方だけ入れ替えると、 TF または FT という別の均衡点になる。つまり FF をデフォルトとすることでフレームの再計算は容易になる。一方、 p や q が観察されたときは矛盾が生じないため、フレームの再計算は起こらない(カードは選択されない)。

このようにデフォルトの変化は、均衡の移動という同一フレーム内の再計算にとどまらず、フレームシステムの再構造化を意味する。次に、否定を結論に導入した場合のゲームモデルを示そう(表6参照)。

表6 ゲーム6のペイオフ表

		C	
		T	F
A	T	0, 0	$\alpha + \varepsilon, 1$
	F	$1, \beta + \delta$	α, β

表6の様式は、表5と同じく行が p に対する真理値、列が q に対する真理値である。 $\alpha, \beta, \varepsilon, \delta$ が、 $0 \leq \alpha, \beta \leq 1, \varepsilon, \delta \geq 0$ の範囲とし、 FF をデフォルトとすると、ゲーム6は p と q をその最適反応において正しく選択させる。このため、ゲーム6を用いて人々の正答率が著しく向上することを説明できる。とくに $\alpha = \beta = 1, \varepsilon = \delta = 0$ のとき、 $p \rightarrow \neg q$ の真理表とペイオフ表の値は一致する。

5. 3 囚人問題

A, B, C の3名がそれぞれ独房に収監され、近々3人まとめて処刑される予定だった。しかし恩赦により1人だけ釈放されることが、なんとクジ引きで、決められた。Aは恩赦の件を知り、看守に自分以外の2人のうち処刑される1人の名前を教えてほしいと頼んだ。看守は翌朝Aに対して「Bは処刑される」と伝えた。このときAは自分の釈放される確率をどう考えたのだろうか？

オリジナルの作者は不明だが、少なくとも Gardner の連載コラム[2][3]まで遡ることができる[14]。その後、1980年代に日本人の研究者によって注目された[6]。数学的に同等の3ドア問題（モンティホール問題）は1990年代に再注目されて、学際的に研究された[14]。

事前確率を3人とも等確率で釈放されるとすると、ベイズの定理を適用して、Aが釈放される事後確率は $1/3$ と計算される。一方、現実の人々が考える直観的な答えの多くは、 $1/2$ である。 $1/3$ で変わらないとするものも少なくないが、ベイズの定理を用いて導出したわけではなく、「AとCは同じ確率のはずだ」「CとAの比は変わらない」「確率は前と変わらない」などと考えた可能性がある（市川と下條は主観的定理と名付けている）。

ところで、もしベイズの定理の答えが正解であるとするなら、囚人Aは、もし確率を正しく算出できれば事前確率とは異なるオッズ比を知ることになるから、ある種の駆け引きを用いて口の堅い看守から情報を引き出すことに成功している。駆け引きは戦略的なゲーム状況として認識されるような、必ずしも協調的でないコミュニケーションである。囚人Aは情報の提供を看守に依頼するにあたって、看守には秘密を守る義務があること、また看守が誠実に答える動機があることを考えたはずである（実際、Gardnerの問題ではコイン投げのメカニズムの発案者は囚人Aである）。看守には黙秘するという選択肢や三味線を弾くという選択肢もありえたと考えれば、問題文章に示された事実として、おそらくAの企てが功を奏した結果、看守は黙秘を選ばなかったのだと解釈することが

可能である。いずれにせよ、Aの情報提供要求に対して回答することで看守は守秘義務に背き情報漏えいしてしまっている。

エージェント自身だけが知る情報のことを私的情報という。（非協力）ゲームを通じて私的情報を自発的に提供させる手法は、メカニズムデザインと呼ばれる。3囚人問題をメカニズムデザインとして読み直すと、囚人Aが看守から自分にとって都合の良い行動を何とかして引き出そうとして、（隠れた）コイン投げを含めたメカニズムを提案し、看守はこれを受け入れて一つのメッセージを選んだということになる。具体的なゲーム設計については次節で詳しく述べるが、釈放確率を推理するフレーム終端を、協調原理にしたがうゲームで埋めると、ベイズ解がとってかわるのを防ぐことになる。

ところで、通常、3囚人問題は確率計算という非戦略的なフレームにあてはめて論じられることが多い。しかし次節で詳述されるように、囚人Aに協力して情報提供する看守の行動は、Grice的なコミュニケーション[4]として再解釈することができ、またゲーム理論はこれを定式化する目的に適する。ちなみに3ドア問題は、確率計算としては3囚人問題と同じ問題であるが、戦略的状況の特色が濃くなるためゲーム理論を適用した文献は複数ある（一例として[8]がある。英語版のWikipediaに表がまとめられている）。

クイズ番組で回答者が暫定的な一つのドアを選び、司会役が一つの外れのドアを開いて見せ、回答者に自分の選択を変える機会を与える。ベイズ解によれば選択変更によって当選確率は倍になる。もちろんこれは胴元側にとって不利であるから、司会者の誠実性の仮定は疑わしい。回答者が直観解に固執することは機会を逃すかわりに騙されるリスクも減じる。また食事問題と呼ばれる別の同型問題は比較的やさしいとされる[6]。食事問題にはコミュニケーションの文脈がなく、戦略的ゲームの状況でもないため、人々は確率の比の計算に集中しやすいと思われる。

6. フレームシステムとしてのゲーム

本節では3囚人問題のゲームモデルを具体的に設計する。ラベリングシステムは均衡点を確認するグラフィカルな補助ツールだが、認知モデルとして用いる場合、ゲーム理論を用いたフレームシステムの表現とみなされる。またコンピュータシミュレーションによるゲーム設計により、直観解の優越を確認する。

Grice [4]の協調原理にしたがうとすると、看守は以下の条件を守る必要がある。

- 聞き手が真の状態について正しく信じているときは黙っている (量の格率)。
- 嘘はつかない (質の格率)。
- 聞き手の信念が誤っていると思えたら訂正する (関係の格率)。
- 簡潔で、曖昧性・多義性がない (様態の格率)。

一方、囚人Aは看守に情報提供を求めているのだから、誰が釈放されることになると、看守にBかCの名前を言ってもらった方が好ましいと思っているはずである。また、もしAが自分自身の釈放を信じているとすると、いずれの名前を聞いても差はないだろう。もしAがBの釈放を信じているならCの処刑を知ったときAは自分自身の信念を確認できる。同様にAはCの釈放を信じているときBの処刑を知って合点がいく。

実際に、看守が協調原則に違反せずに囚人Aに情報提供できるような2人ゲームを設計することができる。

$\Theta = \{\theta_A, \theta_B, \theta_C\}$, $M = \{m_O, m_B, m_C\}$, $B = \Theta$ をそれぞれ状態空間, メッセージ空間, 信念空間とする。真の状態 $\theta \in \Theta$, 看守のメッセージ $m \in M$, 囚人Aの信念ポジション $b \in B$ である。

記号 $\theta_A, \theta_B, \theta_C$ はA, B, Cのうちいずれが釈放されるかという真の状態を区別する。また m_O, m_B, m_C で看守が選ぶメッセージを表す。ただし m_O は黙秘である。囚人Aの選択は自らが信じる真の状態である。これをAの信念ポジションと呼び、真の状態の記号を濫用して表す。看守とAの間の協調的コミュニケーションは表7に示すゲーム7によってモデル化される。

表7 ゲーム7

True state θ_A		Prisoner A's belief		
		θ_A	θ_B	θ_C
warden's message	m_O	1, 0	0, 0	0, 0
	m_B	0, 1	1, 0	0, 1
	m_C	0, 1	0, 1	1, 0

True state θ_B		Prisoner A's belief		
		θ_A	θ_B	θ_C
warden's message	m_O	0, 0	1, 0	0, 0
	m_B	0, 1	0, 0	0, 1
	m_C	1, 1	0, 1	1, 0

True state θ_C		Prisoner A's belief		
		θ_A	θ_B	θ_C
warden's message	m_O	0, 0	0, 0	1, 0
	m_B	1, 1	1, 0	0, 1
	m_C	0, 1	0, 1	0, 0

看守のペイオフ $u_w : B \times M \times \Theta \rightarrow R$ は真の状態に依存すると考えられるため、表7は3つの表からなる。囚人Aは真の状態を知らないので、3表とも囚人Aのペイオフ $u_A : B \times M \times \Theta \rightarrow R$ の値は変わらない。表7中の各セルはペイオフ値の組 (u_w, u_A) を表す。

看守の最適反応は表7の各列の最大値を与える行である (表8参照)。看守の最適反応がつねに黙秘でないかぎり、厳密にいうと守秘義務に反する。そこで「Aが期待していない情報を与えない」ことを弱い意味での守秘義務と解釈する。

表8 ゲーム7における看守の最適反応

True state	Prisoner A's belief		
	θ_A	θ_B	θ_C
θ_A	m_O	m_B	m_C
θ_B	m_C	m_O	m_C
θ_C	m_B	m_B	m_O

表 8 は弱い意味での守秘義務を満たすが、表 8 の列 θ_A において看守のメッセージは 3 状態を識別することに注意されたい。それゆえ、もし看守に対して A が自分の釈放を信じていることにコミットできれば、ゲーム 7 は情報を暴露するプロトコルとして利用できる。実際、信念ポジション θ_A は他の 2 つの信念ポジションを優越 (dominate) するので、看守は表 7 で θ_A の列だけを見てメッセージを選択できる。しかし、現実の人間がつねに合理的な選択の基準にしたがうとは限らない。

ちなみに、3 ドア問題でクイズ回答者は、表 7 と同様のゲームを仮定すると、もし選択が変更されないと司会者に信じさせることができれば自分に有利な情報を引き出すことができる。しかし実際に選択を変えれば動的に矛盾するため、手番のあるゲーム 3 と同じく、均衡点は不完全である。

図 2 はゲーム 7 のラベリングシステムである (ラベル図とも呼ばれる)。図 2 は、上が囚人 A のランダム戦略 p の空間、下が看守のランダム戦略 q の空間を表す。二つの正三角形の高さを 1 とすると、内部の点から各辺に降ろした垂線の長さが確率を表す (つまり $0 \leq p_\theta, q_m \leq 1, \theta \in B, m \in M, \sum p_\theta = \sum q_m = 1$)。各辺では一つの純戦略の確率が 0 であり、外周の弧の交点は原点、弧で仕切られた領域は原点と 3 次元単体が作る四面体の他の面にあたる。

ラベリングシステムは図 2 の各領域に一定のルールでラベルを割り当てる。ラベルはその確率に対する相手エージェントの最適反応および自分自身の確率 0 の純戦略の記号 (ないし番号) を連結したものである。たとえば、直観解のラベルは $m_O m_C \theta_B$ である (図 2 上の三角形の右上上の 3 領域の交点)。

使用されない戦略を無視すると、均衡点では両者はお互いに最適反応を提供し合う。実際、均衡点の必要十分条件は、完全ラベル、すなわち両者のラベルの連結がすべての純戦略をカバーすることである (証明は[15]または[16]を参照)。

真の状態が θ_A のとき看守の各純戦略に対するペイオフの期待値は、

$$v_w(m_O) = p_{\theta_A}, v_w(m_B) = p_{\theta_B}, v_w(m_C) = p_{\theta_C}$$

である。看守の選好を $\succ_w \subseteq M \times M$ (\succ_w は完備的かつ反射的かつ推移的) で表すと、

$$\begin{aligned} m_O \succ_w m_B &\leftrightarrow p_{\theta_A} \geq p_{\theta_B}, \\ m_B \succ_w m_C &\leftrightarrow p_{\theta_B} \geq p_{\theta_C}, \\ m_C \succ_w m_O &\leftrightarrow p_{\theta_C} \geq p_{\theta_A} \end{aligned}$$

という関係が成立つ。三角形内の領域を区分する境界線上ではこれらの不等式が等式で成り立つ。

同様に、囚人 A の選好を $\succ_A \subseteq B \times B$ で表すと、

$$\begin{aligned} \theta_A \succ_A \theta_B &\leftrightarrow q_{m_B} \geq 0, \\ \theta_B \succ_A \theta_C &\leftrightarrow q_{m_C} \geq q_{m_B}, \\ \theta_C \succ_A \theta_A &\leftrightarrow q_{m_C} \leq 0 \end{aligned}$$

となり、ゲーム 7 では 3 頂点を除くと、囚人 A は

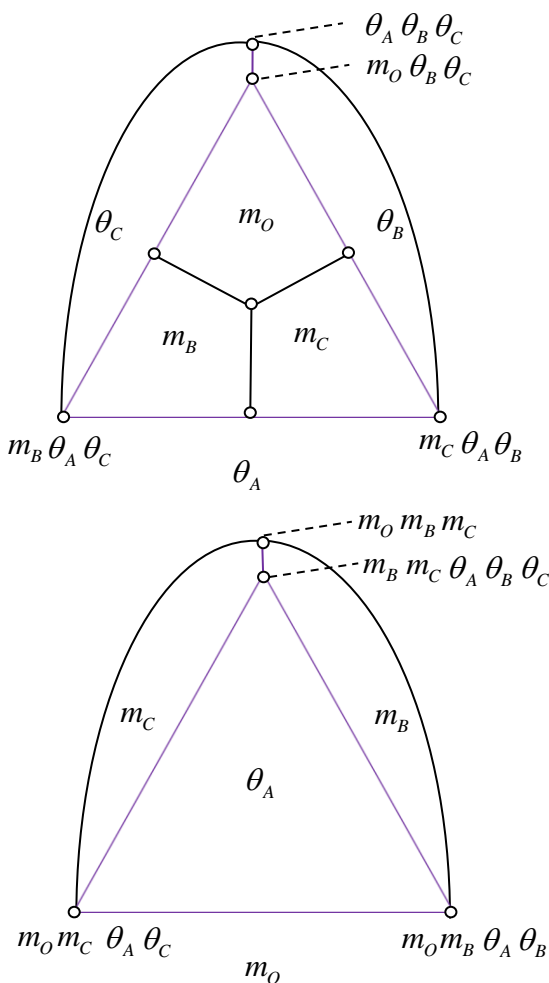


図 2 ゲーム 7 (θ_A) のラベル

つねに自分自身の釈放を信じる。

真の状態が θ_A の場合、ゲーム7の均衡点は4組あるが、看守の最適反応はいずれも黙秘である。したがって看守は厳密に守秘義務を満たすことができる。均衡点における囚人Aの最適反応は、①自分自身の釈放が確実、②1/3の均等確率（ベイズ解の事前確率）、③と④自分自身と他の2人のうち一人がともに確率1/2（つまり直観解）のいずれかである。図2からこれらの確率組はいずれも完全ラベルをもつことが分かる。

3囚人問題を解く思考は、デフォルトの均衡点から別の均衡点への移動として表現される。一つの均衡から離れると不安定になり欠落するラベルを補うように移動するが、この動きが情報反応モデルに対応する。したがって情報反応モデルはゲームの均衡点を求めるアルゴリズムの一般化でもある。たとえば原点の組をデフォルトとしてラベルが一つだけ欠落する道（準相補道）を選ぶと別の均衡点に辿り着く（Lemke-Howson法）。

なお協調原理はつねに守られるべきものというわけではなく、その格率に違反することを通じて会話の含意を派生し、日常のコミュニケーションを円滑にすると考えられている[4]。ラベリングシステムを用いると、直観解あるいはベイズ解が均衡点で信じられるようなゲームを、協調原理にしばられずに、特定できる。そこで最後に、筆者が行ったコンピュータシミュレーションによるゲーム生成実験の結果を紹介する。

ゲームを一定の複雑さの範囲内で設計できるようにするため、以下のように工夫する。まずBとCの名前を交換しても結果が不変となるよう、表9のようにペイオフをパラメータ化する。

表9 パラメータ化されたゲーム

True state θ		Prisoner A's belief		
		θ_A	θ_B	θ_C
warden's message	m_O	α, σ	β, μ	β, μ
	m_B	γ, τ	δ, λ	ϵ, ω
	m_C	γ, τ	ϵ, ω	δ, λ

次にペイオフ値をn水準として真の状態 θ を固定する。表9のパラメータを仮定すると、純粋にベイズ解をモデル化できるゲームは3値以上ではじめて得られる（証明は読者に委ねる）。実際に、0と1の2値のペイオフと-1を加えた3値のペイオフで実験した結果を表10に示す。

表10 ゲーム生成実験の結果†

	2値	3値
直観解あり	181 (111)	1680 (424)
直観解のみ	19 (7)	381 (55)
ベイズ解あり	162 (104)	1620 (468)
ベイズ解のみ	0	321 (99)
両方あり	162 (104)	1299 (369)
Total	900 (384)	22500 (5292)

†括弧内は均等事前確率の均衡点をもつケース

このように、直観解を均衡点で信じるゲームが囚人Aの釈放確率を推理するフレームの端末を埋めると、ベイズ解のゲームがそれにとってかわることはよりむずかしくなると考えられる。

7. まとめ

本論文は思考をゲームとして定式化する方法を提案し、いくつかの例題について具体的にゲームを設計した。情報反応モデルでは、人々が最適反応について直観的に感じとることができることを仮定した。またラベルは、いわば食べ物の養分（＝情報）であり、人々は不足分を補おうとして均衡点に至る。しかし情報反応モデルでは人々がゲームを定式化して均衡点を計算することを仮定しない。ジャンケンで偏った手の出し方をすれば相手に負けることを人々は理解するが、しかしそれは均衡点を計算して等確率になることを計算した結果ではない。なお、互いに口をきけないエージェント（反応ルール）の集まりとして心の活動を表すことは、モジュール形式と呼ばれる哲学的議論にかかわると思われる。本論文のアプローチとの関連付けはオープンクエスチョンである。

参考文献

- [1] Evans, J. S. B. T., and Lynch, J. S., (1973) "Matching bias in the selection task", *British Journal of Psychology*, Vol. 64, No. 3, pp. 391-397.
- [2] Gardner, M., (1959a) "Mathematical games: Problems involving probability and ambiguity", *Scientific American*, Vol. 201, No. 4, pp. 174-184.
- [3] Gardner, M., (1959b) "Mathematical Games: How three modern mathematicians disproved a celebrated conjecture of Leonhard Euler", *Scientific American*, Vol. 201, No. 5, pp. 181-188.
- [4] Grice, H. P., (1975) "Logic and Conversation", P. Cole and J. Morgan (eds.), *Syntax and Semantics*, Vol. 3, Academic Press. Reprinted in Grice, H. P., *Studies in the Way of Words*. Harvard University Press, 1989, pp. 22-40. [清塚邦彦(訳), "論理と会話", 勁草書房, 1998]
- [5] Hurwicz, L., and Reiter, S., (2006) "Designing Economic Mechanisms", Cambridge University Press.
- [6] 市川伸一・下條信輔, (2010) "3囚人問題の展開と意義をふり返って", *認知心理学研究*, Vol. 7, No. 2, pp. 137-145.
- [7] Kirby, K. N., (1994) "Probabilities and utilities of fictional outcomes in Wason's four-card selection task", *Cognition*, Vol. 51, No. 1, pp. 1-28.
- [8] Mueser, P. R. and Granberg, D., (1999) "The Monty Hall Dilemma Revisited: Understanding the Interaction of Problem Definition and Decision Making", University of Missouri. Working Paper 99-06. Retrieved 5 June 2015.
- [9] Minsky, M., (1975) "A Framework for representing knowledge", P. H. Winston (ed.), *The Psychology of Computer Vision*, pp. 211-277, McGraw-Hill. ["知識を表現するための枠組", 白井良明, 杉原厚吉 (訳), コンピュータビジョンの心理, pp. 237-335, 産業図書, 1979]
- [10] Minsky, M., (1986). "Society of Mind", Simon and Schuster. [安西 祐一郎 (訳), "心の社会", 産業図書, 1990]
- [11] Nash, J. F., (1951) "Non-cooperative Games", *Annals of Mathematics*, Vol. 54, No. 2, pp. 286-95.
- [12] Oaksford, M., and Chater, N., (1994) "A rational analysis of the selection task as optimal data selection". *Psychological Review*, Vol. 101, No. 4, pp. 608-631.
- [13] 岡田章, (2011) "ゲーム理論 新版", 有斐閣.
- [14] Rosenhouse, J., (2009) "The Monty Hall Problem", Oxford University Press.
- [15] Shapley, L. S., (1974) "A note on the Lemke-Howson algorithm", *Mathematical Programming Study*, Vol. 1, pp. 175-180.
- [16] Shimojo, Shinsuke, and Shin'ichi Ichikawa, (1989) "Intuitive reasoning about probability: Theoretical and experimental analyses of the "problem of three prisoners", *Cognition*, Vol. 32, No. 1, pp. 1-24.
- [17] Simon, H. A., (1956) "Rational choice and the structure of the environment". *Psychological Review*, Vol. 63, No. 2, pp. 129-138.
- [18] von Neumann, J., & Morgenstern, O., (1947) "Theory of Games and Economic Behavior", Princeton University Press. [銀林浩他(訳), "ゲームと経済行動", 筑摩書房, 2009]
- [19] Wason, P. C., (1966) "Reasoning", B. M. Foss (ed.), *New Horizons in Psychology*, Vol. 1, pp. 135-151. Penguin.
- [20] 山 祐嗣, (2008) "直説法的 Wason 選択課題におけるバイアスをどのように説明できるか?", *認知科学*, Vol. 9, No. 4, pp. 473-486.