

Discussion Paper No.329

ルイス『コンヴェンション』を制度論から読み直す
——前半部の議論を中心として——

中央大学経済学部
瀧澤 弘和

June 2020



INSTITUTE OF ECONOMIC RESEARCH
Chuo University
Tokyo, Japan

ルイス『コンヴェンション』を制度論から読み直す

——前半部の議論を中心として——

中央大学経済学部

瀧澤弘和

2020年5月4日

概要

近年ゲーム理論の一部で、これまでほとんど省みられなかったデイヴィッド・ルイス独自の共通知識概念が新たな関心を引き起こしている。本稿は、この研究潮流の存在を背景として、ルイスの『コンヴェンション』を改めて読み解き、ゲーム理論を経由した現代の経済学的制度論の視点からその射程と限界を見定めることを目的とした研究ノートである。

ルイスの分析枠組は、ある特定の性質を持った事態（基底）が存在することとして共通知識を定義することで、共通知識が生成されるメカニズムを明示的に論じることができる点で独自であり、ゲームをプレーする主体の経験の共有がゲームの均衡プレーを可能にするという「外在主義的」な魅力を持っている。また、コンヴェンションの当事者たちがコンヴェンションに関する知識を持つという主張のように、制度の知識の問題を明示的に取り上げたことは、制度批判や制度変化における反省的思考の役割の考察可能性を開くものである。しかし、ルイスがコーディネーション問題だけに焦点を当てたことは、彼の議論を制度一般の理論へと拡張しようとする際にいくつかの点で慎重でなければならないことを意味している。

また、彼がなぜ「コーディネーション均衡」という一般的でない均衡概念を使用しているのかという、これまであまり明確に回答されていなかった謎について、コンヴェンションが規範の一つの種だとする主張に結びつけた解釈を提示する。

1 はじめに

制度の本質論を語るうえで、デイヴィッド・ルイス (David Lewis, 1941–2001) の『コンヴェンション』(Lewis, 1969) はどうしても避けて通ることができない関門である。このことは、これまで多くの制度論者たちが、たいていの場合には批判的観点からのものであるものの、それを参照点として論じてきたことに表われている。たとえば、ほんのいくつかの例を思いつくままにあげても、Gilbert (1992, 1996)、盛山 (1995)、Sugden (1986)、Gintis (2009)、青木 (2010)、Aoki (2011)、青木 (2014)、Guala (2016)、倉田 (2019) などが、かなりの頁数を割いて真剣にルイスのコンヴェンション論の意義を論じている。

現代の経済学的な制度論は「ゲームのルールとしての制度論」と「ゲームの均衡としての制度論」とに大別されることが多いが、この視点から見た場合には、ルイスのコンヴェンション論は、「ゲームの均衡としての制度論」の原点とも言うべきものである。基本的にルイスはコンヴェンションをゲームの均衡における人々の行動の規則性と捉えているからである。しかし、その後のゲーム理論的・均衡論的な制度論には、主体の限定合理的行動を仮定する進化ゲームの枠組を利用してコンヴェンションを分析するものが登場してきた (Young, 1993, 1998)。実際、「均衡としての制度」の立場に立つ現代の制度論者のなかではどちらかというと、こちらの方が優勢になっているようにも思える。たとえば『比較制度分析のフロンティア』に収められたスカームズ (2016) やサグデン (2016) などの論稿は明確に進化ゲームによる制度の説明で、科学的説明としてはこれで十分だとする立場である。これとの対比から言うと、一見して「伝統的な合理的ゲーム理論」に依拠しているか

のように見えるルイスのコンヴェンション論は、同じくゲームの均衡として制度を捉えるといっても、進化ゲーム理論的アプローチとは対極的な位置を占めている。なぜなら、ある程度まで人間を合理的と仮定しているうえ、制度的現象において人間主体が持つ知識や信念が果たす役割をかなりの程度重視しているからである。

制度論と深く関連しているものの、それとは若干はずれたところに存在するゲーム理論の文脈もある。周知の通り、ルイスの『コンヴェンション』は、今日のゲーム理論において中心的な概念である「共通知識」を最初に提起したのものとして、ゲーム理論の教科書や論文で頻繁に引用されている（たとえば Osborne and Rubinstein, 1994）。また共通知識は、合理的プレーヤーたちが均衡をプレーする際の認識論的条件を与えるうえで中心的な役割を果たすものとして、今日においても、最重要の概念として受け入れられているといえるだろう（たとえば Aumann, 1987; Aumann and Brandenburger, 1995）。しかしながら、こちらの文脈では、『コンヴェンション』が出版されてからしばらく後に Aumann (1976) による集合論的な共通知識の定式化が独立して登場してからというもの、ルイスの定式化は、オーマン以前の原初的な定式化と見なされるようになってしまって、あまり重視されなくなってしまった。いわば、概念を最初に提起した人として、言い訳程度に言及されるような感じになっているのだ。

しかし近年では、ルイスの共通知識概念をオーマン流の集合論的な知識概念に吸収されたものと見なすことの誤りを正そうとする研究が登場してきている。その典型的なものが Cubitt and Sugden (2003) である。これから本論文で詳細に論じることになるが、ルイスの共通知識概念は単に人々の頭のなかにある信念について論じているのではなく、人々が共通に経験する事象を基盤として、どのように共通知識が形成されるのかを説いたものである。このことを改めて評価すべきだというのがキュービットやサグデンらの主張なのである。

ここに、制度論においてルイスが占める位置とゲーム理論の歴史におけるルイスの再評価が交差する余地が生じている。今述べたような、人々が共通に経験する事象をもとにして共通知識が発生し、ゲームの均衡において行動の規則性が生じるという観点から、そのアイデアを改めて現代的な制度論に組み込もうという発想が生じてくるのは、ある意味で自然なことだと言えるだろう。実際、Gintis (2009) や Aoki (2011) はこのような観点から、ルイスの新解釈を彼ら自身の制度論のなかに組み込もうとした試みと見なすことができる^{*1}。

本稿は、近年のゲーム理論における再評価と、ルイスのコンヴェンション論が制度論のなかで占める位置の同定のために、まずはルイスの原著に立ち戻って、そのコンヴェンション論の理路を明らかにしようという試みである。

本稿の構成は以下の通りである。まず、第2節において『コンヴェンション』の全体の章立て、全体的な概要を簡単に抑えておきたい。そのうえで、本稿の焦点が前半のコンヴェンション一般の議論にあることを確認する。第3節では、『コンヴェンション』第I章の内容に沿って、ルイスの問題設定の仕方を確認することにする。コーディネーション問題、コーディネーション均衡、コンヴェンションとそれが成立するための十分条件としての整合的な相互的諸期待の体系に関する議論がここに含まれる。第4節は、第II章の前半部分で展開されている議論のロジックをやや詳細に追っていく。そこで中心的な役割を果たすのがルイス独自の共通知識概念であり、それをういたコンヴェンション概念の再定義である。ここでの議論は「インディケートする」等、独自の概念がふんだんに使用されたもので、おそらく、その晦渋さから多くの読者を遠ざけてきたものと思われる。その論理展開のプロセスをできるだけ明快に示したい。ここが本稿の山場ということになろう。第5節は、共通知識概念を用いたコンヴェンションの再定義に焦点をあて、その含意としてのコンヴェンションそのものに対する集団の知識の問題について述べる。第6節では、コンヴェンションの規範性に関するルイスの議論を吟味する。第7節で総括する。

*1 私自身の知る限りでは、後年ルイス自身が最初の自分の最初の著書についてどのように感じていたのかについてはあまり直接的な証拠はない。Guala (2016) は第7章へのノートのなかで、ルイスが「共通知識」という言葉を用いたことを後悔し、「共通信念」と呼ぶべきと考えていたことに言及している。ルイス自身によるものとしては、Lewis (1975) があるが、これはもっぱら言語のコンヴェンションに関して、その誤りを正そうとするものであり、コンヴェンション論一般に関しては実質的な変更を提案していない。ルイス自身が、自分の著作に対して、今の制度論が問題にしているような観点から評価していたかどうかは疑わしい。

2 ルイス『コンヴェンション』の概要

周知の通り、デイヴィッド・ルイスは20世紀後半を代表する分析哲学者であり、様相実在論ないし可能世界論における、かなり極端な立場の主張者としてよく知られている(たとえば(野上, 2020)を見よ)。ここで取り上げようとしている『コンヴェンション』は、彼の分析哲学者としてのキャリアの出発点近くに位置するもので、ハーバード大学でW・V・O・クワインの指導のもとに執筆した博士論文を出版したものである。したがって、当然のことながら、その執筆動機、推論過程、結論は当時の分析哲学的背景を有している。飯田(1995)によれば、1960年代には様相論理の意味論をはじめとする非標準論理の研究が盛んになると同時に、それを用いて、フレーゲ以来展開してきた形式言語の論理を自然言語に適用する新たな機運が生じてきた時期であった。ルイスの後半の言語に関する議論には、この潮流に一石を投じようとする強い意欲が感じられる。

しかし、本書執筆の少なくとも表向き^ルの動機は「言語はコンヴェンションに統治されているという常識的な考え方」(Lewis, 1969, p.1)に独自の考察を加えて、最終的にはこの見方を擁護しようとするのであった。当時は、「論理的・数学的真理は規約(convention)による」という規約主義的テーゼがカルナップたちによって提出されたことに対して、クワインが反駁を加え、それが決定的な反駁であると思われていた時期である。それと同時に、分析性の概念を可能世界論に基礎づけるという目論みも本書の執筆動機の一つであった。このことは、クワインが本書に寄せた序文、そして本書の最終部分に見ることができる。

まず本書の構成について確認しておこう。

序文(クワインによるもの)

イントロダクション

I コーディネーションとコンヴェンション

- 1 コーディネーション問題の実例
- 2 コーディネーション問題の分析
- 3 コーディネーション問題を解決する
- 4 コンヴェンション
- 5 コンヴェンションの実例

II コンヴェンションを精緻化する

- 1 共通知識
- 2 コンヴェンションの知識
- 3 コンヴェンションに対するオルタナティブ
- 4 コンヴェンションの度合い
- 5 コンヴェンションの帰結

III コンヴェンションと他の諸概念との比較対照

- 1 合意
- 2 社会契約
- 3 規範
- 4 ルール
- 5 同調的行動
- 6 模倣

IV コンヴェンションとコミュニケーション

- 1 シグナルの実例
- 2 シグナリングの分析
- 3 言葉シグナリング

4 シグナルのコンヴェンションの意味

5 シグナルの意味_{nm}

V 言語のコンヴェンション

1 諸可能言語

2 文法

3 可能言語における意味論

4 真実性のコンヴェンション

5 集団における意味論

全体はかなり明確に大きく二分することができる。第 III 章までの前半部が一般的なコンヴェンション論であり、後半の第 IV 章と第 V 章がそのコンヴェンション概念を言語に対して適用して論じた部分である。本書執筆の動機から言って、実際には後半部が重要ということになるが、後半部の議論が成功したかどうかはさておいても、前半部において展開されたコンヴェンション一般に関する議論はそれだけでも十分な価値を持っている*2。

ところで、ここで「コンヴェンション」という言葉についても少し述べておきたい。コンヴェンションという言葉はラテン語の *convenio*(不定形 *convenire*) という動詞に由来するものである。これは「ともに」を表わす *con-*と「来る」を表わす *venio*(不定形 *venire*) が結合したもので、そのコアにあるイメージは「集合すること」と推測されるが、多くの場合に「合意する」ことを意味する。今日でも人が多く集まる集会をコンヴェンションと言い、そのような場所をコンヴェンション・センターとかコンヴェンション・ホールと呼んでいるが、コンヴェンションという言葉には人が集まるということが含意されているのだろう。上述したように、カルナップたちの立場(コンヴェンショナリズム)を「規約主義」と訳す場合には、この合意するというニュアンスが込められているように思われる。また、本書の序文の冒頭でクワインが、「子供の頃のわたしは、レンブラント風に、テーブル沿いに座っている特別評議員会の厳粛な会議 (*convention*) においてわれわれの言語が決定され、通達されるものと思い描いていた」と述べるときのコンヴェンションもそのようなニュアンスを帯びている。

「集合する/合意する」といった意味の方だけから考えると、「言語はコンヴェンションによる」というテーゼを擁護することは、一見してきわめて荒唐無稽なことに思われるかもしれない。しかし英語の *convention* には他方で、慣習とか伝統といった意味もあるのである。私の知る限りでは、デイヴィッド・ヒュームがコンヴェンションという言葉が持つ、こうした多義性を活用した最初の哲学者と言えるかもしれない。Hume (1740/2012, 第 3 巻) には「コンヴェンションに入る (*enter into*)」という使用法があり、参考文献に挙げた邦訳ではコンヴェンションを「合意」と訳しているのだが、確かに「合意」と理解してじっくりいくところが多い。しかし、ヒュームは言語や貨幣もコンヴェンションであると指摘していたし、社会契約論に対する批判的見方からしても文字通りの合意が必要だと考えていたか、あるいはコンヴェンションの持続にとっては合意に対する記憶が重要であると考えていたかという、そうとは言えないだろう。

ルイスはこうした多義性を利用しつつ、彼自身が本質的と考える要素—その一部はヒュームも述べているような、共通利益の感覚と、相手の行為を参照して自分の行為を行うという相互依存的意思決定の状況であろう—を抽出して、彼独自のコンヴェンションの概念化を行うのである。もちろん、ルイスのコンヴェンション概念においても合意の要素がまったく排除されているわけではなく、合意がコンヴェンションの形成に大き

*2 クワインは当然、ルイスが出した結論には納得できなかった。しかし、序文のなかで次のように述べている。「分析的真理と総合的真理を区別するという問題は、明らかにこの研究の動機の一つであった。ルイスは最終的に、コンヴェンションの概念はこの区別の核心でないと結論づける。しかし彼は、この理由から分析性の概念を受容不可能だと見なすわけではない。彼はむしろ、ある人たちが始めたところに行き着き、分析性の概念を可能世界の概念に基礎づけるのである。分析性問題のこのような決着に彼が満足していることは、結局のところ、それが彼のコンヴェンション研究の動機の大きな部分となりえたのはどうしてかと考えさせるものである。しかし、われわれは彼がどのような動機を持っていたにしろ、それに対して感謝してよいだろう。というのは、読者はこの本の中身で、分析性ではないにしても、言語哲学における鍵概念としてのコンヴェンションを理解するようになるからである」。

な役割を果たすことは明示的に述べられている。

かくして、ルイスが本書で俎上に載せようとしているコンヴェンションは、ひとまずは、規約、合意、慣習などが無いまぜになった多義的な言葉であることに留意する必要がある。本書を『社会的慣習』とか『規約』のように訳すことが難しい理由はここにある。ルイスは彼がコンヴェンション概念の中核にあると考えるものを抽出して論じたうえで、それを合意、社会契約、規範等々の周辺概念と比較してもいる。これが第 III 章のテーマである。

本書の後半では、前半で精緻化した一般的なコンヴェンション概念を言語に適用する。最初に、意味を前提としないコミュニケーションの状況を定式化して、シグナリング問題、シグナリング・システム、シグナリング・コンヴェンションを次々と定義し、シグナリングが言葉によって行われる言葉シグナリング・コンヴェンションにおける意味論を考察する。ここまでが、第 IV 章である。直説法的な文に関するルイスの意味論はタルススキの意味論に近いものである。しかし、命令法の文に関する意味論の展開には、ルイスのオリジナリティが感じられる。

第 V 章では、より複雑化した「可能言語」という概念を一般的に定義したうえで、この言語のあり方を分析する。最終的にルイスが主張したいことは、「集団 P のメンバーたちがある与えられた可能言語 \mathcal{L} を使用している際の言語のコンヴェンションは、 \mathcal{L} における真実性のコンヴェンションとして、もっともよく記述できるだろう」ということである。言語的現象をより現実近づければ近づけるほど、複雑化が生じてくる。言語のなかで発話された文の意味は、発話機会の文脈にも依存することを認めざるをえない。こうして、文の意味を定義するなかで、可能世界における発話機会でその文が真であるか否かまで考えざるをえなくなるのである。

ひとたび可能世界を含めて文の真偽を問うというパースペクティブが得られると、分析的な真理は、すべての可能世界において真であることとして定義される。これが、本書の背後にある執筆動機として、クワインが言及した部分に関連する。しかし、哲学プロパーの議論は私の専門外なので、後半部分の議論の当否やその後の受容のされ方などについては専門的に語るができない。本論文の焦点はあくまで、一般的なコンヴェンション概念を精緻化していく部分、しかもその中心にある共通知識概念を精緻化していく部分に当てられることになる。

3 コーディネーション問題とコンヴェンション

3.1 コーディネーション問題

ルイスのコンヴェンション論でもっとも重要なポイントの一つは、彼がコンヴェンションの概念を一貫してコーディネート問題のみにかかわらせ、コーディネート問題の解決策としていることである。すぐこの後に見るように、コーディネート問題とは、コーディネート・ゲームにおける複数均衡の問題である。ルイスが本書を書いた後に、経済学においてゲーム理論が幅広く応用されるようになり、さまざまなタイプのゲームの分析が多くの知見を生み出してきたことを知るわれわれにとって、このことはきわめて制約的に見えるかもしれない。しかしそこには、彼自身が述べているように Schelling (1960) のゲームに対する見方—ゲーム一般を純粋利害対立から純粋コーディネートへと至るスペクトラムのなかに位置づける—からの強い影響がある。この論点—とりわけ、ルイスのコーディネート問題解決の提案が幅広い適用可能性を持つにもかかわらず、なぜコーディネート問題にこだわったのかという問題—は後で再び取り上げることにし、まずは彼が問題を設定する仕方に沿って見ていくことにしよう。

ルイスはまず第 I 章 1 節において 11 個のコーディネートの実例をあげたうえで、それらに共通する特徴を「コーディネート問題」として抽出していく。今日の標準的な言葉を使用して言うならば、コーディネート問題は基本的にコーディネート・ゲームの状況である (ゲーム的状況のことをルイスは一貫して「相互依存的意思決定の状況」と呼んでいる)。また、そのなかでも特に複数均衡を持つものに限定する。コーディネート・ゲームのもっとも極端なケースは、戦略の組み合わせのすべてでプレーヤーたちが同じ

利得を得る純粋コーディネーション・ゲームであるが、ルイスが考察対象としているのは、純粋コーディネーションのケースに限定されるわけではなく、それに近いものもコーディネーション問題としている。また、ルイスが用いる均衡概念はプロパー・コーディネーション均衡と呼ばれるものである。コーディネーション均衡の概念はのちほど重要になってくるのだが、ナッシュ均衡とそれを精緻化した標準的な均衡概念に慣れているわれわれにはわかりにくいものである。そこでルイスの記法を若干離れて、今日の標準的な記法を用いて、ルイスのコーディネーション均衡の定義を説明しよう。

今日の標準的なゲーム理論のテキストでは、 n 人の非協力ゲームは通常、次の 3 組として定義される： $G = (N, \{A_i\}_{i \in N}, \{g_i\}_{i \in N})$ 。ここで $N = \{1, \dots, n\}$ はプレーヤー集合である。そのジェネリックな要素を i と書くことにしよう。各プレーヤー i は自分の戦略集合 A_i を持ち、そのなかから自分の行為を一つ選択する。そのジェネリックな要素 (1 つの行為) は小文字で a_i と記すことにする。すなわち、 $a_i \in A_i$ である。ゲームとは、(たいていの場合) 2 人以上のプレーヤーたちが関与しているので、その結果はプレーヤーたちが選択した行為の組み合わせによって決まってくる。そこで、改めてこの組み合わせを $a = (a_1, \dots, a_n)$ と書くことにする (通常、戦略結合ないし戦略プロファイルと呼ばれているものである)。 $a \in A \equiv \prod_{i=1}^n A_i$ である。各プレーヤーの利得は、すべての戦略プロファイルの集合 (各プレーヤーの戦略集合の直積集合) から実数への関数である。すなわち、 $g_i : A \rightarrow \mathbb{R}$ であり、 $a \mapsto g_i(a)$ である。戦略プロファイルは、プレーヤー i とその他に分けて、 $a = (a_i, a_{-i})$ と書くことと便利である。他のプレーヤーたちの行為選択を固定しておいて、あるプレーヤー i の行為だけを変化させるという操作が頻繁に行われるからである。ここで a_{-i} は i 以外の戦略のプロファイルである。

さて、ルイスの分析もまた、通常のゲーム理論と同様にナッシュ均衡から始まっている (ただし、彼は単に「均衡」と呼んでいる)。(純粋戦略の) ナッシュ均衡は次のように定義される。 a^* がナッシュ均衡であるのは、以下の条件が成立するとき、そのときに限る。

$$\forall i \forall a_i \quad g_i(a^*) \geq g_i(a_i, a_{-i}^*)$$

この条件は、どのプレーヤーも自分だけがそこから逸脱して、均衡における利得を上回ることができないということを意味している。これに対して、ルイスが本書のなかで独自に提案しているのが「コーディネーション均衡」というもので、次のように定義される。 a^* がコーディネーション均衡であるのは、

$$\forall i \forall j \forall a_j \quad g_i(a^*) \geq g_i(a_j, a_{-j}^*)$$

が成立するときである。その意味は、誰であっても、自分あるいは他のプレーヤーの誰か一人が逸脱したときに、均衡利得を上回ることがないということである*3。

さらに、今日であれば通常、厳密ナッシュ均衡と呼ばれるであろうものは、本書においてはプロパー均衡と名づけられている。すなわち、 a^* がプロパー均衡であるのは、以下の条件が成立するとき、そのときに限る。

$$\forall i \forall a_i \quad (a_i \neq a_i^* \rightarrow g_i(a^*) > g_i(a_i, a_{-i}^*)).$$

また同様にして、 a^* がプロパー・コーディネーション均衡であるのは、

$$\forall i \forall j \forall a_j \quad (a_j \neq a_j^* \rightarrow g_i(a^*) > g_i(a_j, a_{-j}^*)).$$

たとえば、表 1 において、 $(R1, C1), (R2, C2), (R3, C3)$ はコーディネーション均衡であり、 $(R4, C4)$ はコーディネーション均衡でないナッシュ均衡である。もちろん、定義により、コーディネーション均衡はナッシュ均衡である。

ここでルイスがなぜ、ゲーム理論で標準的に用いられるナッシュ均衡ではなく、わざわざコーディネーション均衡という概念に訴えなければならなかったのかは、多くの人にとって謎であるに違いないが、筆者の知る

*3 ルイス自身の言葉では次のように定義されている。「コーディネーション均衡を、一人の主体だけ—彼自身かほかの誰か—が別様に行為したとしても、誰もよくなるような組み合わせとして定義する」(p.14).

		列選択者			
		C1	C2	C3	C4
行選択者	R1	1, 1	0, 0	0, 0	0, 0.5
	R2	0, 0	1, 1	0, 0	0, 0.5
	R3	0, 0	0, 0	1, 1	0, 0.5
	R4	0.5, 0	0.5, 0	0.5, 0	0.2, 0.2

表1 コーディネーション均衡

限り、これに対する回答はルイス自身も述べていないし、ルイスについて論じた論文や書籍においても、解釈はほとんど与えられていない。より詳しくは第6節において述べるが、私の推測は、コンヴェンションの規範性を論じたかったからというものである。

さらに、コーディネーション問題が一意的均衡しか持たないケースを排除するというのも重要である。ルイスは、2人・有限・純粋コーディネーション・ゲームについて、もしそれが一意の均衡しか持たないならば、「支配された戦略」が存在することを証明している。支配された戦略を消去したゲームも同じように2人・有限・純粋コーディネーション・ゲームなので、消去を繰り返していけば、ゲームは1つの戦略プロファイルしか持たない自明なものへと退化してしまう。そのため、一意の均衡しかもたないコーディネーション・ゲームは、コーディネーションという観点からすると自明であるとして、このケースを排除するのである。やや細かいことだが、ここで注意しておきたいのは、ルイスがここで言っている「支配された戦略」が標準的ゲーム理論で言われている「支配された戦略」とは異なり、多くの場合に、NWBR(決して弱い意味でも最適反応にならない戦略)と呼ばれているものに対応していることである。

以上の考察は次のように簡潔にまとめられている。

要約すると、コーディネーション問題—わたしのあげた11の例がお互いに類似しあっている重要な点において、これらの例に似ている状況—とは、2人かそれ以上の主体による相互依存的な意思決定の状況であり、利害の一致が支配的で、2つ以上のプロパー・コーディネーション均衡が存在しているようなものである(p.14)。

3.2 コーディネーション問題の解決としてのコンヴェンション

われわれは日常的にコーディネーション問題を解決している。しかし、ひとたび理論的に問題を設定してみると、コーディネーション問題がそれほど容易に解決されるとは思えなくなってくる。これは、われわれが日常的に協力しあっていることを、囚人のジレンマという理論モデルを用いて説明することの難しさに類比的である。

コーディネーションの実現を説明しようとするときに問題となるのは、主体たちがどのようにして複数ある均衡のなかからどれか一つのものへと自分たちの行為をコーディネートすることができるのかという点である。これはゲーム理論において「均衡の精緻化」と呼ばれてきた問題に近いように思われるかもしれない。一般に、ゲームは少しでも複雑になると複数均衡が出現してしまう。そのなかから、「もっともらしい」一つのを理論的に選択しようとするプログラムが均衡精緻化である。1980年前後に盛んに行われたこのプログラムでは、多くの場合、合理性に対する仮定を強化することで、問題を解決しようとしていた*4。ルイスもまた、ある程度の論理的・認知的能力をもった合理的主体を対象としているので、この点でもルイスの解決しようとする問題と均衡精緻化とが類似しているように思えるかもしれない。

*4 もちろん異論もありうる。むしろ完全均衡のように、ゲームのプレーに摂動を導入して均衡精緻化していくアプローチは一種の限定合理性と見なされることもあった。また、1980年代後半に入ると、明示的に限定合理的な主体を想定して均衡精緻化に取り組む試みがさまざまになされることになる。

しかし、ルイスが考えようとしている問題は、通常ゲーム理論家が考えてきた問題とは次の点で明確に異なるものである。第一に、均衡精緻化の理論家たちが、ゲームが与えられたときに、それを合理的にプレーする仕方が一意的に決まるはずだという暗黙の前提のもとに精緻化のプログラムを推進していたのに対して、ルイスが考えるコンヴェンションにとって複数均衡の存在は必須の条件となり、しかも、それらの均衡の間で人々が基本的に無差別であることを重要視していることである。いわばゲームの複数均衡性を自分自身の概念化に積極的に取り入れているといえる。第二に、ルイスは、主体たちが一回限りのコーディネーション・ゲームをプレーしているとも、完全に同一のコーディネーション・ゲームに毎回直面しているとも想定しているのではなく、一連の類似した問題に直面する中で一定の行為選択のパターンが持続していくのはどうしてなのかという問題を考えていることである。ここには、二つの論点が含まれている。第一には、実際には正確に同一な問題など一つとしてない多様な状況をどのようにして類似的な問題として認識していくのかということであり、第二に、一連の状況に直面するとき、毎回新たなコーディネーション・ゲームがプレーされるときとでは、どのように異なるメカニズムが作用しているのかという問題である。数学的に定式化されたゲームの枠組に留まり、それだけを資源としながら、主体がどのように思考し行動するのかを考える通常のゲーム理論的アプローチよりも少し幅を拡げて、ゲーム理論を道具として使用しながら、人間のコーディネーション問題に迫ろうとしていると表現してもよいかもしれない*5。

さて、本論に戻って、われわれがどのようにしてコーディネーションを達成するのかという問題を考えてみよう。もちろん、主体たちがランダムに選択したとしてもコーディネーションに成功する可能性は存在する。しかし、「適切な仕方で整合的な相互的諸期待の体系を媒介することで、成功する可能性が高くなる」(p.25)とルイスは言う。この表現で注意しなければならないのは、整合的な相互的諸期待の体系を媒介することはコーディネーション問題を解決するための「一つ」の方法と見なされているのであって、必要条件と見なされているわけではないということである。

本書のなかでは主体はある程度は合理的であると見なされており、基本的に信念と選好の組み合わせによって、行為を選択するという古典的な実践的推論が想定されている。図1は、「あなたとわたし」が待ち合わせをするとき(第I章の冒頭に掲げられている11の例の(1)である)に作用する実践的推論の状況を図示したものである。「わたしは、あなたがそこに行くだろうと期待する」という「信念」と「あなたがそこに行くという条件のもとで、わたしはそこに行くことを望む」という選好とが組み合わさって、「わたしはそこに行くことを望む理由を持つ」ということが「含意」される。ここまでが実践的推論の部分であると言ってよいだろう。これに対して、この結論から「わたしはそこに行く」というところに伸びている図中の細い線は「合理的主体の心的状態または行為の間の因果的結合」(p.28)を示している、とルイスは述べる。

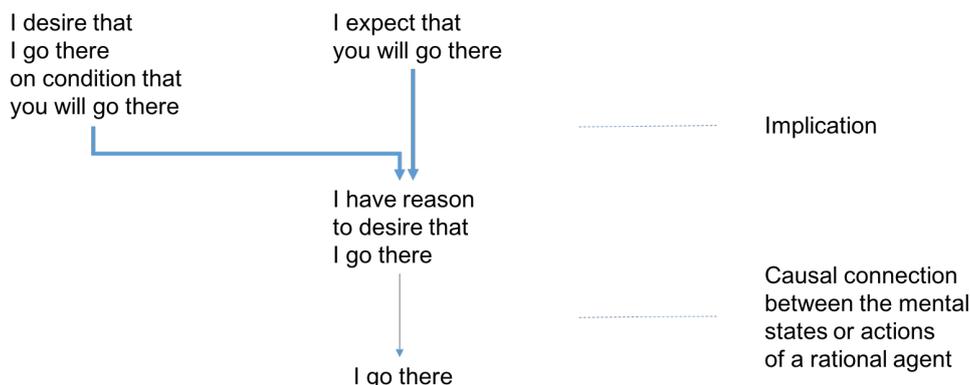


図1 わたしの基本的な実践的推論 (p.29)

しかし、この推論の前提である「わたしは、あなたがそこに行くだろうと期待する」という信念は、どこ

*5 通常のゲーム理論はこのような観点からは「内在主義」的と特徴づけられるかもしれない。

からやってくるのか。ルイスによれば、それは、わたしがわたしの思考のなかで、あなたの推論を「複製 (replacate)」することによって得られるものである。その様子を図示したのが図2である。ここでは先の図では前提とされていたものが結論となっている。この思考プロセスはすべてわたしの思考のなかで行われているとはいえ、あなたの推論を辿ることによって得られたものである。さらに、あなたの推論の複製だけではなく、ここでは先の図になかった合理性に関する期待が登場している。すなわち、「～する理由を持つ」から「～する」への移行において、その人の合理性が関与することが明示されているのである。相手の合理性に対する期待は、この後「補助的前提」(の一つ)として明示化される。もちろん、このプロセスはさらに続けることができよう。中央上部の「わたしがそこに行くだろうと、あなたが期待することを、わたしは期待する」を導くために、わたしの思考のなかのあなたは、わたしの思考を複製することによって、この結論を導くことだろう。こうして、わたしの思考のなかで、わたしはあなたの思考を複製するが、今度はその先で、あなたはわたしの思考を複製している・・・とように続いていく。

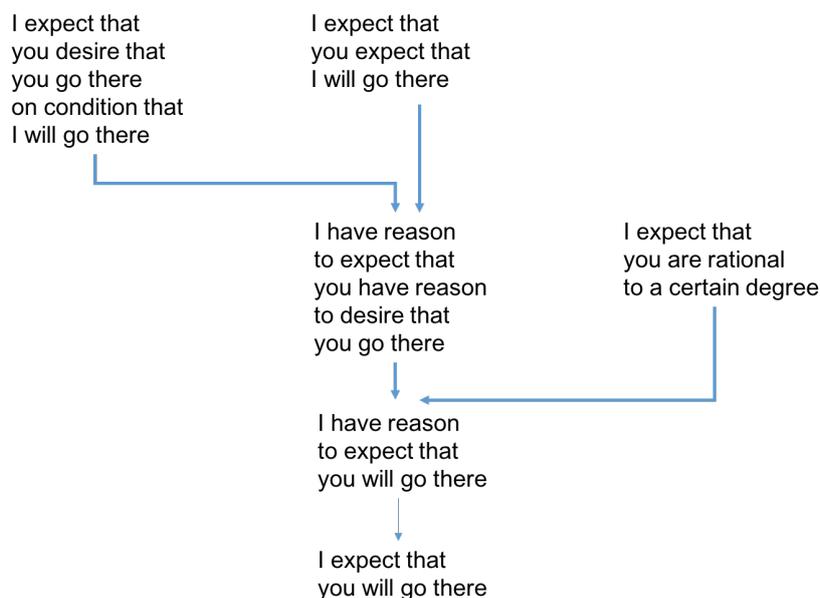


図2 相手の期待を入れ子状に持つ (p.29)

このツリーはわたしの思考とその帰結を表しているが、当然、あなたの思考とその帰結がもう一方の側に存在している。そして、両方の期待の体系が互いにマッチしているときに、均衡における戦略の組み合わせが実現することになる。それが「整合的な相互的諸期待の体系」と呼ばれる状態である。この概念は誤解されやすいと思われるので、いくつかの点をクラリファイしておきたい。

第一に、ここで論じているような主体たちの期待における複製は、現実世界におけるインタラクションの結果として行われ、期待を整合的なものにするわけではないということである。このことをルイスは次のような言葉で表現しているが、その真意は明確であろう。

複製は、人々の間を行ったり来たりするインタラクションではないことに注意されたい。それは、一人の人が世界に関する自分の信念の帰結を導くプロセスである。この世界は、自分たちの信念の帰結を導く人々を含み、その中では、他の人々に対する彼らの信念が含まれる・・・と彼が信じているものである。世界におけるわれわれのインタラクションによって、われわれは、前提として役立ちうる、さまざまな高階の期待を獲得する。その後のわれわれの推論においては、われわれはお互いを写し出し、お互いを写し出すお互いを写し出す・・・ためにベストを尽している窓のないモナドである (p.32)。

第二に、ここでの議論はややもすると、こうした整合的な相互的諸期待の体系が「必要」と述べてい

るように取られてしまいがちであるが、ルイスはそう言っていないことである。高階の期待の成立は必要条件ではなく、あくまで十分条件である。そして、ここでは「何らかの仕方」で高階の期待が手に入るかもしれないと述べているのみである。再びルイス自身の言葉を聞いてみよう。

必要とされる補助的諸前提を所与として、行為に関する任意の高階の期待は、主体に対して、行為の選択に関する一つの十分に良い理由を与える。この一つの理由は、彼の選択を全体として十分に正当化する多数の理由のなかの一つの理由かもしれないし、通常はそうであろう。ある主体が何らかの仕方、自分とパートナーの行為に関するいくつかの階の期待を持つことがもともと正当だと仮定しよう。そして、補助的諸前提が利用可能であると仮定しよう。このとき、彼のもともとの諸期待のそれぞれは、あれやこれやの仕方で行う理由を独立に与えることになる。もし彼が幸運ならば、これらの理由のすべてが同一の行為に対する理由となるだろう。このとき、その行為は、余分な仕方では正当化されるので強く正当化される。彼はそれを行うことに対して、彼のもともとの期待のうちのいずれか一つだけによって提供されたよりも多くの理由(単数)を持っている(pp.32-33)。

ここで、当然、どのようにして高階の期待の整合性が実現するのかという疑問が湧くことになる。この問題は第II章において本格的に取り組まれる課題となるのだが、第I章でも簡単に論じられている。ルイスは整合的な相互的諸期待を生み出す手段として、合意(明示的合意、約束、現在の意図の宣言)、顕著さ(salience)、そして先例をあげている。顕著さは「ある目立った点において一意性を持ち、それによって残りのものから際立っていること」(p.35)として定義され、この性質に注目したSchelling(1960)の業績が言及される。そしてまた、先例もまた顕著さの一例であると論じられるのだが、これこそがこれから主として論じていく論点にもっとも関連するものである。すなわち、先例はコーディネーション均衡をプレーするプロセスを駆動し、それを持続させていく。このプロセスがもっとも適切な意味でコンヴェンション的なのである*6。

どうして先例はとりわけ強力なコーディネーションの源泉となるのか。一言でいえば、以前に成功して行われたことが類似の状況でも適用可能だと考える人間の傾向性があるからである。このとき、われわれはその傾向性を互いに期待し、そのことを期待し・・・という形で高階の期待を持つかもしれない。

先例というメカニズムによるコーディネーション均衡の達成では、先例をつくりあげた主体と後からそれに従う主体の一致すら必要とされない。また、以前のケースで、なぜコーディネーションが達成されたのかということも問題にならない。そして、それはその後も持続する可能性が大きい。現時点で先例に従うことによって、われわれはそれ以降に利用可能な先例のストックを追加していることになるからである。

さらに、多数の先例が蓄積して、それが曖昧さもなく利用可能になるときは、われわれは正確に同じ先例を知っていることも必要ないし、その知識が詳細なものであることも必要でなくなる。後に共通知識がどのように成立するのかを論じることになるが、先例に対する知識が正確に同じでなくても構わないという点は、共通知識の維持にとって非常に大きなメリットとなるだろう。

たとえば、日本の道路で自動車が左側通行をしているのを頻繁に見ていれば、そのことから、日本のドライバーたちがほとんど常に左側通行すると期待することができるだろう。また、わたしが運転するときに出会うドライバーたちも同じような経験を通して同じような期待を抱き、彼らもわたしが同じように期待していることを期待しているだろうと推測できる。次節の内容を先取りすることになるが、共通知識が生成される際には、特別な性質を持った事態—共通知識の基底と呼ばれる—の存在が必要とされる。今述べたドライバーのケースにそのことを適用するならば、このケースは、ある事態が基底として機能する際に、必ずしもすべての人たちが正確に同一の事態を見ている必要がないことを示唆している。人々の間に共有経験があることで、それが共

*6 第III章の合意の項を見よ。そこでは、行動の規則性が合意によって立ち上げられ、同意に拘束されているうちは真性のコンヴェンションではないと考えているふしが伺われる。たとえば、以下のように言っている。「実際、わたしの定義によれば、合意によって開始されたコンヴェンションは、合意の直接的効果が薄れる時間を経過して始めて、コンヴェンションとなるのかもしれない」(p.84)。

通知識の生成に役立つのである*7。これは、われわれの日常的なコーディネーション問題解決の説明として、きわめて直観的ではないだろうか*8。

ルイスは第 I 章の最後に、ここまでの議論を総括して、暫定的なコンヴェンションの定義を与えている。

定義 1 (コンヴェンションの最初の定義*9)。集団 P のメンバーたちが再起的状況 S における主体であるとき、彼らの行動における規則性 R がコンヴェンションであるのは、 P のメンバー間の S のどの事例においても、以下のことが成立するとき、そのときに限る。

- (1) 全員が R に同調する
- (2) 全員が、他の全員が R に同調することを期待する
- (3) 全員が、他の全員がそうするという条件のもとに、 R に同調することを選好する。というのは、 S はコーディネーション問題であり、 R への一様の同調は S におけるプロパー・コーディネーション均衡だからである。

ここでは、コンヴェンションが「行動の規則性」として定義されている。しかし、もちろん単なる行動の規則性ではなく、集団 P のメンバーたちが期待や選好に基づいて同調していることが重要である。さらに、コーディネーション問題を考えているので、複数均衡があることが前提とされていることにも注意が必要である。この定義は第 II 章において、共通知識の定義を経由することによって修正される。

4 共通知識

4.1 整合的な相互的期待を創出するものとしての共通知識

これまでのところで、何らかの仕方では、ある程度の高階の整合的な相互的期待が形成され、それによって合理的な主体同士のコーディネーションが可能になることが説明された。そして、高階の期待を形成するメカニズムの候補として、合意、顕著さ、先例があげられた。「コンヴェンションを精緻化する」というタイトルを持つ第 II 章において、ルイスは「事態 (state of affairs)」、 「信じる理由を持つ」、 「インディケートする」、 「帰納的標準」、 「背景情報」、 「合理性」といった概念を駆使して、高階の整合的な相互的期待がどのように形成されるのかをより詳細に分析し、コンヴェンションの定義を修正している。

ルイスがそうしているように、例を使用することで基本的なアイデアを理解するのが近道だろう。前節でも出てきた待ち合せのコーディネーション問題を例にとる。ちなみに、これは合意による高階の整合的な相互的期待形成の例になっている。

次のような事態 A が成立しているとしよう。「あなたとわたしが会って一緒に話をしているのだが、用事が済む前に、あなたが退席しなければならないとする。そこであなたは、明日、同じ場所に戻ってくると言う」(p.52)。この事態 A が高階の期待を生み出してくれる。「明らかに、わたしはあなたが戻ってくると期待するだろう。あなたは、あなたが戻ってくことを、わたしが期待していると期待するだろう。あなたが戻ってくことをわたしが期待していることを、あなたが期待していることを、わたしは期待するだろう。もしかしたら、さらに 1 つか 2 つの階層が存在しているかもしれない」(ibid.)。どうしてか。

事態 A は次の 3 つの条件を満たしている。

*7 ルイスの議論は共通知識の生成メカニズムの理解には役立つが、誤解してはならないのは、それが実際のコンヴェンションの起源とは異なることである。ルイスの議論の焦点は、コンヴェンションがいかんして開始されるか、ないし生成されるかというよりも、それがいかんして維持されるのかということにあるといえよう。ただし、第 III 章 1 節「合意」において、「明示的合意がコンヴェンションのいくつかの可能的な起源の一つにすぎないことを認めたいので、なお、それが特別の地位を享受するのだろうかという疑問が残るかもしれない、コンヴェンションはすべて、合意によって開始することもできたろうというのは、もしかしたら真なのだろうか」(p.86) と自らに問いかけたうえで、これに反論を加えている。

*8 このことは本書において、ここ以外にはそれほど深く分析されていないが、Cubitt and Sugden (2003) では、「信じるための分散した理由 (distributed reason to believe)」というタイトルのもとに、そのロジックが詳細に分析されている。

*9 Lewis (1969, p.42)

- (1) あなたとわたしは、 A が成立していると信じる理由を持つ。
- (2) A は、わたしたちの両方に対して、あなたとわたしが A が成立していると信じる理由を持つことをインディケートしている。
- (3) A は、わたしたちの両方に対して、あなたが戻るだろうことをインディケートしている。

この「インディケートする」とはどういうことか。事態 A が主体 x に対して、命題 p をインディケートするとは、 A が成立すると信じる理由を x が持っているならば、それによって (thereby), x は p であることを信じる理由を持つことをいう。なかなかわかりにくい表現なので、Cubitt and Sugden (2003) の記法を用いて、この定義を頭に入れることから始めよう。

まず、「主体 x が命題 p を信じる理由を持つ」ことを、 $R_x(p)$ と書くことにする。次に「事態 A が主体 x に対して、命題 p をインディケートする」を $A \text{ ind}_x p$ と表わす。そのうえで、上の定義をパラフレーズするならば、以下のようなになる。

$$A \text{ ind}_x p \iff R_x(A \text{ holds}) \rightarrow R_x(p)$$

ただし念のために付け加えておくと、右に表われる論理式で \rightarrow は、通常の論理学で用いられる実質含意を表わすものと文字通りに受け取ってはならない。後で触れるように、ここには帰納的推論が含まれているからである。「それによって (thereby)」という言葉にはそんなニュアンスが込められている。

こうした記号を用いて、上の3つの条件を書き換えれば(ここでは簡単化のために、「われわれの両方に対して」を、添字 w で表わしている),

- (1) $R_w(A \text{ holds})$
- (2) $R_w(A \text{ holds}) \rightarrow R_w(R_w(A \text{ holds})) \iff A \text{ ind}_w R_w(A \text{ holds})$
- (3) $R_w(A \text{ holds}) \rightarrow R_w(\text{you will return}) \iff A \text{ ind}_w (\text{you will return})$

となる。ここで(3)を(2)に適用する。(2)の右辺の内側の $R_w(A \text{ holds})$ に対して、(3)の「含意」をモドウス・ポネンスのように適用することで(われわれが信じる理由を持つ一階の階層のなかで),

$$(4) R_w(A \text{ holds}) \rightarrow R_w(R_w(\text{you will return}))$$

が得られるが、これは「 A はあなたとわたしの両方に対して、あなたが戻るだろうと両方が信じる理由を持つことをインディケートしている」すなわち、 $A \text{ ind}_w R_w(\text{you will return})$ が成立していることになる。上の(3)は $A \text{ ind}_w (\text{you will return})$ であったから、一段高い階の「信じる理由」が形成されているのがわかるだろう。さらに同じ仕方で(4)と(2)を使えば、

$$(5) R_w(A \text{ holds}) \rightarrow R_w(R_w(R_w(\text{you will return})))$$

が得られる。こうして「信じる理由を持つ」が何重にも重なった高階の文が次々と生成されるわけである。

ここで、ある事態が何かを「インディケートする」際には、合理性、帰納的標準、背景情報が含まれているとルイスが述べている点に注意しておきたい。しかし、ここでのルイスの説明はきわめてわかりにくいものである。そこで、Cubitt and Sugden (2003) が挙げているわかりやすい例を使用して説明しよう。彼らは次のような例をあげている。正常な視覚を持つ二人の個人を i と j とする。今、 i と j が同じ部屋にいて、その部屋が稲妻に照らし出された。これを事態 A とする。そして、 p を「数秒内に、雷がなるだろう」という命題とする。このとき、「 A が i と j の両者に、 p が真であることをインディケートする」ということは、すべての正常な大人に支持されるような帰納的推論であろう。インディケートとは、このようにさまざまな経験的知識を駆使することを含んだ概念である。

さきほどの分析では説明を簡略化するために、「われわれ」という「主体」を登場させ $R_w(p)$ のような記号を用いたが、実際に存在するのは「あなた」と「わたし」である。このときには、帰納的標準と背景情報が両者の間で共有されているという仮定がより明示的な役割を果たすことになる。

A が x に対して何かをインディケートするならば、そして y が x の帰納的標準と背景情報を共有しているならば、 A が y に対しても同じことをインディケートしなければならないと考えてみよう。したがって、もし A が x に対し、 A が成立していると信じる理由を y が持っていることをインディケートしているならば、そして A が x に対してを _____ をインディケートし、 y が x の帰納的標準と背景情報を共有していると信じる理由を x が持っているならば、 A は x に対して、 y が _____ を信じる理由を持っていることをインディケートする (その理由は、 A が成立していると信じる理由を y が持っていることである)(p.53).

ここでは、以下のような推論をしていると思われる。ルイスが用いている _____ の代わりに、命題を p と書こう。上の引用部分の「したがって」以下の最初の仮定は

$$R_x(A \text{ holds}) \rightarrow R_x(R_y(A \text{ holds}))$$

と書ける。二番目の仮定は

$$R_x(A \text{ holds}) \rightarrow R_x(p)$$

ここで y も共通の帰納的標準と背景情報を持っていると仮定すれば、

$$R_y(A \text{ holds}) \rightarrow R_y(p)$$

したがって、

$$R_x(A \text{ holds}) \rightarrow R_x(R_y(p))$$

が得られる。共通の帰納的標準と背景情報を用いて、次々と高階の期待が形成されていく様子を示したのが図3である。ここでは、上が低階で下に行くほど高階になっていることに注意されたい。

事態 A が持っている性質に戻ろう。これまでは (1) 式 $R_w(A \text{ holds})$ を使用してこなかったが、これまでしてきたような仕方でインディケートするをパラフレーズして書き下すならば、前件には常に同じ表現 $R_w(A \text{ holds})$ があるので、ここでも推論規則モドゥス・ポネンスを適用することができる (ルイスはこの操作を「分離の原理」を使用すると表現している)。まず (1) を (3) に適用すれば、

(3') われわれのどちらも、あなたが戻ることを信じる理由を持つ

が導かれ、(1) を (4) に適用すれば、

(4') われわれのどちらも、あなたが戻ることを信じる理由を相手が持つことを、信じる理由を持つ

と次々に高階の期待が導出される。

ここから先は、「信じる理由を持つ」から「信じる (期待する)」への転換である。この転換には、合理性に関する補助的前提が必要で、「何かを信じる理由を持つものは誰でも、十分な程度の合理性を持っているならば、そのことを信じるようになるだろう」(p.55) とルイスは述べている。たとえば、(3') において、もし両者が十分に合理的であれば (つまり、0 階のレベルで合理的であれば)、

(3'') われわれのどちらも、あなたが戻ることを期待する

に転換できる。(4') に対して、相手の合理性に対する 1 階の「信じる理由」(すなわち、自分は相手が合理的であると信じる理由を持つ) を適用するならば、「われわれのどちらも、あなたが戻ることを相手が期待していることを、信じる理由を持つ」となり、これに対して、0 階のレベルの合理性を仮定することで、「われわれのどちらも、あなたが戻ることを相手が期待していることを期待する」となる。このように、この合理性の要請が高階の「信じる理由を持つ」に適用されるときには、それに応じて 1 階下のレベルでの合理性が必要とされる。

ルイスは、ここで生成されたような文の無限の連鎖を「これは含意の連鎖であって、誰かの実際の推論のステップではない」と言う。したがって、ルイスによれば、この連鎖が無限の長さを持つこと自体に不適切なこ

I believe that you share my inductive standards and background information

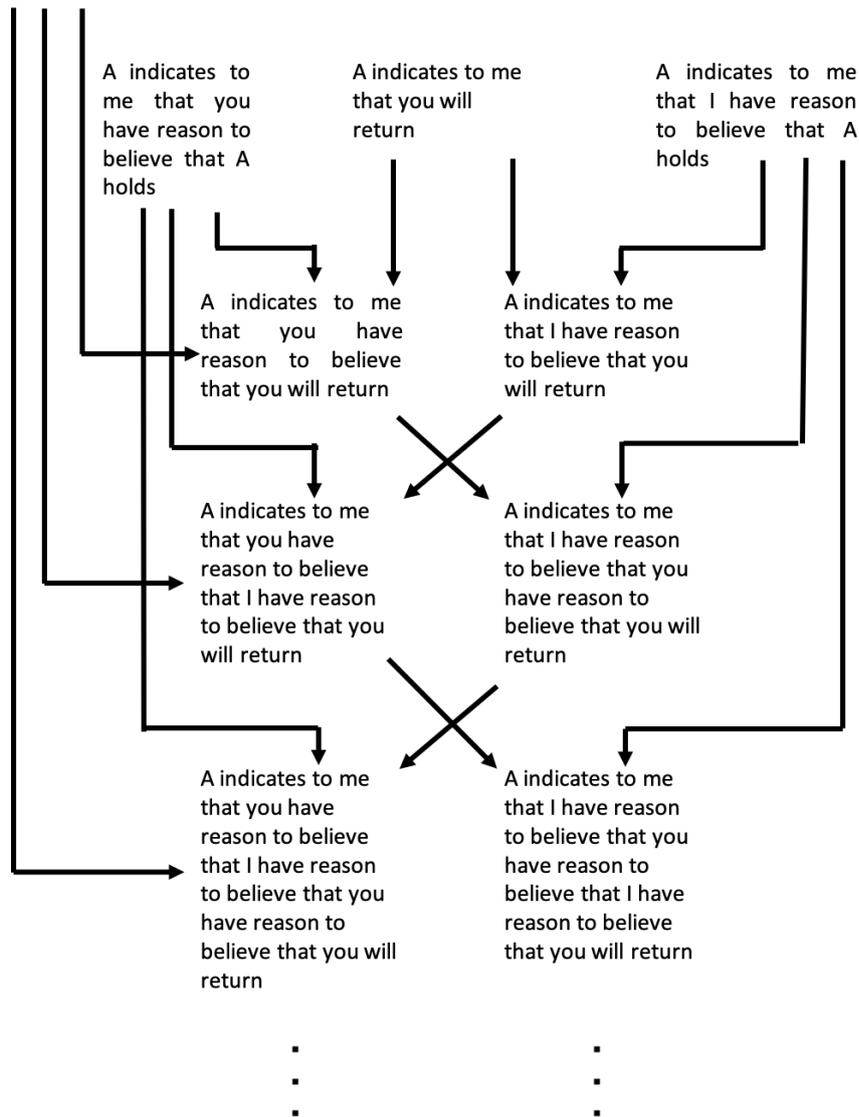


図3 高階の期待の形成 (p.54)

とは何もない。この無限の連鎖を長くしていけるのは、合理性に対する期待の仮定を適用することによってである。しかし高階の期待の形成を保証する合理性の度合いは「明らかに急速に増大する」(p.56)。「このことが最初の二三の階層の期待だけが形成される理由である。この生成プロセスは、補助的な前提が尽きるときに停止するのである」(ibid.)。このステップが合理性という補助的前提だけでなく、主体が持つ帰納的標準と背景情報の共有のあり方にも依存していることは明らかである。

以上で、ある性質をもった事態が存在することが高階の期待を生み出す可能性の分析が終わったことになる。以上の分析を一般的に概念化したものが「共通知識」である。この概念は次のように定義される。

定義 2 (共通知識^{*10})。_____であることが集団 P において共通知識であるのは、次のような条件を満たす事態 A が成立しているとき、そのときに限る。

^{*10} Ibid., p.56

- (1) P のなかの全員が、 A が成立していることを信じる理由を持っている。
- (2) A は、 P のなかの全員に対して、 P のなかの全員が A が成立していることを信じる理由を持っていることをインディケートしている。
- (3) A は、 P のなかの全員に対して、_____ であることをインディケートしている。

このとき、事態 A を P における _____ の基底と呼ぶ。

この段階で、前節において整合的な高階の相互的諸期待という概念が導入されたときの議論と本節の共通知識の議論とを結合して、ルイスがこれらの概念がどのようにコーディネーションの成功に寄与することをイメージしていたのかを確認しておく必要がある。

前節で強調しておいたように、整合的な高階の相互的諸期待の体系が成立することは、決してコーディネーションが実現するための必要条件として導入されたわけではなく、十分条件として導入されたのであった。すなわち、ある程度の高階の整合的な相互的諸期待が成立しているならば、低階へと降りていくことが可能で、最終的にコーディネーション均衡がプレーされると議論されたのだった。しかし、その時点では、どうしてそのような高階の整合的な相互的諸期待が成立するのかについては、合意、顕著さ、先例というメカニズムがあげられながらも、「世界におけるわれわれのインタラクションによって」とか「何らかの仕方です」という表現で述べられていたにすぎない。

本節で導入された共通知識の概念によって、この問に対する回答が与えられたことになる。それはある事態 A が基底となることで、「潜在的に」高階の期待が生み出されるということである。潜在的というのは、高階の期待を生み出すステップは通常、主体同士の間で相手の合理性をどの程度の高階のレベルで帰属させていくのかということと、共通の帰納的標準と背景情報の相互的帰属を仮定できるのかということに依存しているからである。このようにして、共通知識の基底をもとにして高階へと上昇していった期待が、コーディネーションの具体的行為に向かって降りていくという構図が描かれるのである。

4.2 ルイス共通知識論の意義

先に述べたように、ルイスの共通知識の概念は、その後 Aumann (1976) が数学的にきわめて洗練された共通知識の概念化を提出したことによって、その前段階的な概念と見なされて、その内容が省みられなくなっていった。現在の標準的な共通知識の理解について簡単に整理しておこう。インフォーマルには、ある事象 E が主体 1 と主体 2 の間で共通知識であることは、単に 1 と 2 が E を知っているだけのこと（これは「相互知識」と呼ばれている）ではなく、1 が E を知っていることを 2 が知っていること、2 が E を知っていることを 1 が知っていること、1 が E を知っていることを 2 が知っていることを 1 が知っていること、2 が E を知っていることを 1 が知っていることを 2 が知っていること・・・がすべて成立していることをいう。すなわち、 $K_i(E)$ で主体 $i(= 1, 2)$ が事象 E を知っていることを表わすとすると、

$$K_1(E), K_2(E), K_1(K_2(E)), K_2(K_1(E)), \dots$$

がすべて成立することである。

オーマンはこれを次のように数学的に定式化した。世界状態すべての集合 (状態空間) を Ω とする。 $\omega \in \Omega$ が発生したとき、主体 i はこれをピンポイントで認識できるわけではなく、集合 Ω 上に分割 \mathcal{P}_i を持ち、 ω を要素として含むこの分割の要素 $P_i \in \mathcal{P}_i$ が生じていることだけしか知ることができない。この分割は情報分割と呼ばれているが、より細かいものから、粗いものへと順序づけることができることは直観的に理解できるだろう^{*11}。より細かい情報分割の方がより細かい情報が表現できることがわかるだろう。1 と 2 の分割のどち

^{*11} フォーマルに言うと、分割とは Ω の非空の部分集合の集合で、その任意の 2 つの要素 (それ自身 Ω の部分集合) が互いに素で、すべての要素の和集合が Ω に一致するものである。分割 \mathcal{P}'_1 が \mathcal{P}_1 より (弱い意味で) 粗いとは、任意の \mathcal{P}'_1 の要素が \mathcal{P}_1 の要素の和集合で表現できることをいう。

らに対してもより粗い共通の分割のなかでもっとも細かいものを最密共通疎化 (finest common coarsening) と呼ぶ。ωにおいて事象 E が 1 と 2 の間で共通知識であるとは、E が ω を含むような最密共通疎化の要素を包含することをいう。こうした定式化が、多数の直観に反する結論を含みながらも、複数の主体の間に発生するインタラクティブな認識を厳密に扱うことを可能にしてきたことは間違いない (分かりやすい導入としては Osborne and Rubinstein, 1994, を見よ)。

純粹に形式的なオーマンの理論的枠組に対して、ルイスの共通知識は共通知識の基底と呼ばれる特殊な性質を持った事態の存在によって定義され、主体の帰納的標準の共有によって実質化される。そうすることで、主体が共有した経験がどのように高階の期待を形成しうるのかを説明する道が開かれている。もちろん、オーマン流の定式化に沿いながら、ルイスのアイデアを展開することも可能である*12。しかし、オーマンの枠組では、主体は自分が知らないことを知っているという「知恵の公理」など、実際のわれわれの知識について不自然なことが、ほぼ前提に近いところに置かれてしまっている。もちろん、帰納的推論が果たす役割を論じる余地もほとんどない。制度論的な観点からすると、われわれが外的世界に生み出しているさまざまな記号が共有され、共有された体験を生み出すことによって、われわれのコーディネーションを可能にしていると考えられることはもっともらしいことであり、ルイスのアイデアが制度的現象の説明に利用できる可能性が高いというわけである。

また、整合的な相互的期待の体系がいかにか創出されるのかを論じるなかでルイスが行っていることは、実質的にはゲームの均衡 (ナッシュ均衡) がどのようにしてプレーされるのかの説明になっていることにも注意しておきたい。実際、ギンタスは次のように言っている (以下の引用文では、本稿の用語に沿って訳文を変更している)。

ルイスが、先例に基づくコンヴェンションに関する基本的説明を純粹コーディネーション問題のみに適用しようと考えたことには驚かされる。実際には、均衡が厳密ナッシュ均衡になるという事実以外には、ルイスは純粹コーディネーション・ゲームの性質をほとんど使っていないからだ (ナッシュの名前は索引にない)。インフォーマルには、コーディネーション問題のそれ以外の性質についてもルイスは議論している。一人のプレイヤーの戦略の変更がすべてのプレイヤーにとって損になること、プレイヤーたちは複数の均衡のうちのどれが実際に選ばれるかという点に関してほとんど無関心であることといった点である。だが、私の見る限り、複数の厳密ナッシュ均衡をもつゲームからなるはるかに幅広いゲームの集合に対してルイスの考えを適用することができる。ルイスは「利益の実質的対立」が存在しないことが彼の分析にとって重要だと考えていた。だが、実際には、そのような仮定が彼の結論を導くためにフォーマルに利用されている箇所は一つもない (Gintis, 2009, 邦訳 p.242)*13。

しかし上述したように共通知識の概念化が異なれば、その結果として、たとえば Aumann and Brandenburger (1995) におけるようなナッシュ均衡の認識論的条件とはまったく異なる結論が得られるのも当然と言えよう。

オーマンとブランデンバーガーの主要定理によれば、すべてのプレイヤーたちが共通事前分布を共有しており、合理性が相互知識であり、プレイヤーたちの信念の組み合わせが共通知識であることが、その信念の組み合わせが (混合戦略) ナッシュ均衡となることの十分条件だというものである*14。ここでは余分な詳細には立ち入らないが、彼らのステートメントにおける「プレイヤーたちの信念が共通知識である」という件は、文字通りに、どこまでも高階の信念が成立することを意味している。これに対して、ここでのルイスの解決策は、プレイヤーたちの経験の共有が「基底」となって、お互いのプレーに対するある程度まで高階の整合的な相互的期待が形成され、そこからプレイヤーたちは期待を低階まで下げていくことによって、均衡プレーの実現に

*12 たとえば、Vanderschraaf (1998) はオーマンに近い枠組でルイスの共通知識を再構成する試みである。

*13 私もこの意見に部分的に同意するが、ギンタスもまた、なぜルイスがナッシュ均衡でなく、コーディネーション均衡を使用していたのかという問題を見逃しているように思う。

*14 共通事前分布の共有とは、世界状態の集合に関する確率分布を完全に共有していることを意味している。またここでの相互知識とは、お互いに相手が合理的だという知識を持っているということである。もちろん、モデルを大幅に変更もしない限り、必要条件について語ることは意味をなさない。ランダムにプレーしてもナッシュ均衡が実現することがあるからである。

成功しようというものである。頭の中の世界での超合理的な信念形成の世界と、現実世界における経験共有に依存した信念形成との違いと言ってもよいだろう。ここに近年の制度論で出てくる主体たちを取り巻く記号システムが均衡プレーを媒介するという考え方の親和性の鍵がある。

4.3 ルイスにおける正当化と現実

しかし、ルイスはわれわれが行為を選択する際に現実に起こっていることそれ自体をリアルに記述しようとしていたのだと考えることは誤りである。ルイスはわれわれがコンヴェンションに従って同調行動をとるときに、あまり考えることなく単なる習慣でそうする可能性を認めているし、毎回、整合的な相互的期待を考えぬいて行為を選択しているとは考えていない。たとえば、第IV章の「シグナリングの分析」においては、次のように言っている。この引用は、シグナリング・システムがコンヴェンションを構成することを論じていく途上における記述である。

心ここにあらずで、シグナリングが単調でルーチン化している主体のことを考えてみよう。その人は立ち止まって、以下のように考えるだろうか。「この行為は、受け手をかくかくしかじかに行動させることになるだろうし、それが目下の状況では最善だから、わたしはこの行為をした方がよい」(主体がメッセージの受け手の場合)、あるいは「メッセージの送り手がそうしているのは、かくかくしかじかを観察したからで、その場合には、わたしにとってこれが最善の行為となるだろう。だから、わたしはこの行為をした方がよい」(主体が受け手のメンバーの場合)。答えはノーである。その人はリラックスして、習慣によって自分の状況依存的計画に従っているのだ。その主体は、相手に従っている計画を所与として、現在従っている計画が最善であることに一般的な仕方では気づいている。これに反するような証拠を受け取ったときには、停止して思考し、異なる状況依存的計画に従って行為し始めるだろう。しかし、物事がうまく運ばれている限り、自分がしていることがどうして特定の機会にベストとなるのかについて考えたりしないのである。(p.141)。

多くの場合に習慣的に行為しているのだとすれば、主体たちが整合的な相互的期待を形成して行為しているという表現は、どのような意味を持つのだろうか。ルイスは上の引用に続く一節において、これに対して回答している。

しかし、われわれには主体たちの実際の推論を表現する必要がない。われわれが考慮しなければならないのは、信念と欲求を所与にしたときに主体たちがたどることができるだろうような実践的推論によって、彼らの選択を合理的に正当化することだけである。しかしこのことは、彼らの選択を説明することに対する関心を放棄していることを意味しない。正当化は、主体がその正当化に従った推論プロセスを実際にたどるか否かに関わりなく、選択を説明するのである。というのは、その正当化を考え抜かないときでも、われわれが信念と欲求によって正当化される仕方で行為する傾向にあることは、人間本性の事実だからである。このことは否定的な仕方でもよいだろう。われわれの選択を実際にコントロールしている習慣的プロセスがどのようなものであったとしても、そのプロセスがわれわれの信念と欲求に反する傾向を持ち始めるならば、ただちに明示的な実践的推論によって上書きされ、修正され、再訓練されるであろう (ibid.)。

ルイスが自分の議論を「正当化のモデル」のようなものとして考えていたことには、主体にある程度の合理性を要請することとも関係がある。正当化の背景を何ら持たない子供たちは、コンヴェンションの集団に属していると見なされていないからである。われわれ人間は、正当化を背景として行為している。その正当化の背景は、共通知識の基底のような「事態」の経験的な共有に基づいているものの、集団のメンバー全員で同じものである必要はない。集団は共有された経験によって、コーディネーションを可能にする、多様な正当化の背景を資源として持っているという描像が描けるのではないだろうか。

普段われわれは正当化をする必要もなく、習慣的にコンヴェンションに従っていたとしても、それが何らかの変化に晒されるとき(ルイスは、「そのプロセスが信念と欲求に反する傾向を持つとき」と言っているが)、その正当化を可能にするような明示的・論理的な階層が決定的に重要になってくる。しかしこのときには、われわれは正当化の知識に基づいた討議を行うのだろう。正当化のレイヤーとそれに対する知識は、このような意味で重要なのである。後にわれわれは、われわれがコンヴェンションに同調するとき、コンヴェンションに対する知識を有していることについての検討がなされることになる。しかし、先回りをしすぎてしまったので、コンヴェンションの議論に戻ることにしよう。

5 コンヴェンション

5.1 共通知識概念を用いたコンヴェンションの再定義

共通知識の定義は一般の命題 p に対して与えられているので、これをコーディネーション問題を解決する整合的な相互的諸期待の体系を生み出すさまざまな要素に対しても適用することができるだろう。前小節において例証されたのは、「あなたが戻るだろう」という「信念」に関する高階の期待の形成が、ある事態の存在によって可能になることであった。信念に関する高階の期待形成が可能なのであれば、それだけでなく、選好に関する高階の整合的な相互的諸期待や、合理性に関する高階の整合的な諸期待も、何らかの事態の存在に支えられた共通知識によって生成されうるのではないだろうか。ルイスはこの問いに対して、イエスと回答する。たとえば、さきほど待ち合わせの例(これはさきほども注意したように「合意」の事例である)のなかで使用した事態 A のなかで、お互いに、相手がもとの場所に戻るならば、自分もそうしたいという選好が表明されているならば、それは選好に関する共通知識を形成することになるだろう。

そこで、ルイスは前章で定義したコンヴェンションの定義の修正を提案する。先の定義のなかで、コンヴェンションの成立条件としてあげられていた項目が、集団のなかで共通知識でなければならないと要請するのである。こうして修正されたコンヴェンションの定義は以下のようなものである。

定義 3 (修正されたコンヴェンションの定義^{*15})。集団 P のメンバーたちが再起的状況 S における主体であるとき、彼らの行動における規則性 R がコンヴェンションであるとは、以下のことが真であり、 P において共通知識であるとき、そのときに限る。すなわち、 P のメンバーにおける S にかなる事例においても

- (1) 全員が R に同調する。
- (2) 全員が R に同調することを期待している。
- (3) 全員が、他の全員がそうするという条件のもとに、 R に同調することを選好する。 S がコーディネーション問題であり、 R への一様な同調が S におけるコーディネーション均衡だからである。

コンヴェンションの定義において共通知識を要請することについて、ルイスは二つの理由をあげている。第一には、「単純に、われわれの例に共通する重要な特徴のすべてを定義に書き入れたいからであり、関連する事実の共通知識がそのような特徴の一つであると思われるからである」(p.59)。第二には、共通知識を要請しなければ、コンヴェンションに入れたくないような状況もコンヴェンションにされてしまうからである。たとえば、自動車運転の例で、人々が「自分以外の誰もが、何の理由もなしに、習慣によって右側通行している。皆、他の人たちの行動についてどんな期待を抱いていたとしても、右側通行を続けることだろう」(ibid.) という誤った信念にもとづいて左側通行をしている状況もコンヴェンションになる。ルイスはこういう状況を排除したいのである。

共通知識であることを要請するということは、ある性質をもった事態 A が存在しなければならないと要請することである。前章では、合意、顕著さ、そして先例が整合的な相互的諸期待の体系を生成すると考えられて

^{*15} Lewis (1969, p.58)

いたが、共通知識の概念を設定した今、問題はこれらの概念がどのようにして共通知識の基底たりうるのかという間に置き換えられることになる。そして、顕著さはより弱い基底にしかかなりえないものの、合意と先例は一般的に十分協力的な共通知識の基底として機能することが論じられる。

5.2 コンヴェンションの知識について

第II章2節以下では、第1節で得られたコンヴェンション概念の含意とさらなる展開に当てられている。ここではそれらすべてを振り返るのではなく、第2節で議論されている「コンヴェンションのについての知識」を取り上げたい。

ここでのルイスの主張は、もし集団 P において、ある事態 B が成立していることが共通知識になっているのであれば、 P のなかの誰であっても、 B が成立していることが共通知識であると期待する理由を持つということである。したがって、集団 P におけるコンヴェンションがあるとき、そのコンヴェンションの定義に出てきた各項目が共通知識になっているのであれば、集団 P の全員はそのことを信じる理由を持つと論じられる。もちろん、合理性を仮定すれば、「信じる理由を持つ」は「期待する」ないし「信じる」と置き換えられるだろう。

ここでの論証はフォローすることが難しいものだが、私自身が誤解しているリスクを承知のうえで解釈を与えてみたい。共通知識の定義によって、ある事態 B が成立していることが集団 P において共通知識になっているのであれば、基底となる事態 A が存在し、以下のことが成立している。

- (1) P のなかの全員が、 A が成立していることを信じる理由を持つ。
- (2) A は、 P において、 P のなかの全員が、 A が成立していることを信じる理由を持つことをインディケートしている。
- (3) A は、 P において、 B が成立していることをインディケートしている。

これらは、次のようにパラフレーズすることができる。

- (1) $R_P(A \text{ holds})$
- (2) $R_P(A \text{ holds}) \rightarrow R_P(R_P(A \text{ holds}))$
- (3) $R_P(A \text{ holds}) \rightarrow R_P(B \text{ holds})$

(1) と (2) からモドウス・ポネンスにより、

- (4) $R_P(R_P(A \text{ holds}))$

が導かれ、(2) と (3) のそれぞれを、その内容を信じる理由を持つと書き直すことができると考えれば(ルイスはそのように言っているように思える)、次の (5) と (6) が得られる。

- (5) $R_P(R_P(A \text{ holds}) \rightarrow R_P(R_P(A \text{ holds})))$
- (6) $R_P(R_P(A \text{ holds}) \rightarrow R_P(B \text{ holds}))$

(4) から (6) を見てみると、以前の (1) から (3) をすべて R_P で囲んでしまったことがわかる。(1) から (3) は「ある事態 B が成立していることが共通知識である」という内容だったので、(4) から (6) によって、「そのことを P のなかの全員が期待する理由を持つ」と言っているように思われる。さらにルイスは「(4), (5), (6) から、われわれは P のなかの全員が、条件 (1), (2), (3) を満たす事態 A が存在すると信じる理由を持っていると、推論してよいであろう」(p.61) と言う。したがって、

特にコンヴェンションが共通知識の一項目として成立しているならば、そのコンヴェンションが成立している集団に属していること—その一部であること—は、ある意味で、それが成立していることを知ることである。規則性 R が集団 P におけるコンヴェンションならば・・・ P における全員が、 R がコンヴェンションであることを信じる理由を持っていないければならないということである (p.61)。

ただし、この知識は集団のなかで経験的に習得されるものであって、不可謬なものでも完全なものでもないといふのは言う。一部の制度論者が言うように、制度的知識がア・プリオリなものだとはしないわけである*16。

また、われわれのコンヴェンションに対する知識は実際には、きわめて貧弱なものであるかもしれない。ルイスがあげる理由は三つある。第一に、われわれは自分たちのコンヴェンションがコンヴェンションの定義条件を満たしているという結論へと導く証拠を持っていたとしても、その結論へと至る推論を完遂していないかもしれない。第二に、コンヴェンションはヒュームのボートの漕ぎ手たちのように非言語的なものであるかもしれない。第三に、われわれのコンヴェンションに対する知識は、一度にひとつずつ取られた特定事例に制限された知識であるかもしれない。確かに規則性がコンヴェンションとなる際には「一般的な」期待が存在しているのだが、その一般性は合成的な意味で (in sensu composito) ではなく、分割の意味で (in sensu diviso) 一般的であるにすぎないかもしれない*17。

コンヴェンションに関する知識についてのルイスの議論はきわめて興味深い。この観点こそ、制度とは何かを語る際に欠かすことのできない再帰性 (reflexivity) を制度論に取り入れる道を開くものかもしれない。あるコンヴェンション的な行動の規則性に同調する集団に属しているとき、われわれはそのことを少なくとも潜在的には知ることができる。冒頭でも述べたように、今日では進化ゲーム理論を用いた制度の概念化が優勢になっているように思われるが、そうしたモデル化においては、主体たちは自分たちの行動に対する知識を持たず、環境に適応して行動を変化させるだけなので、制度変化はあくまで外生的な変化によって与えられる以外にない。しかし、われわれはわれわれが一定のコンヴェンションに同調していることを知っているがゆえに、制度変化や制度のあるべき姿について公共討議の場で論じることができるのである。これは青木昌彦が初めて彼独自の制度の概念化を提起して以来、一貫して重視していた観点でもある (Aoki, 2001)。

ルイスはコンヴェンションに対する知識に関連して、スタンリー・キャヴェルの主張を取り上げている。キャヴェルによれば、ネイティブ・スピーカーは、自分が言うだろうことについて言っていること (that he says about what he would say) に関して、正当化の証拠を必要としないという。ルイスはこの発言を「ある人をひとたび、われわれの言語のネイティブ・スピーカーとして認めたならば、その人がわれわれのコンヴェンションの当事者であることをすでに認めたことになる。したがって、その人はコンヴェンションが規定することを知っているのだ」(p.62) と解釈する。このように言うとき、ルイスはコンヴェンションの当事者として認めることの意味について論じているのだろう。

ここでのルイスの議論もそうであるが、知っていることを知っているという記述が論理学の言語のなかに取り入れられたとしても、その言語を十分に拡張しなければ、知っていることを批判するというを形式的に表現することは困難である*18。しかし、キャヴェルの議論に対するルイスの解釈は、形式化された論理学の言語の範囲をすでにある程度逸脱して、集団のメンバーシップに関する含意について述べているように思われる。このような拡大解釈が許容されるのであれば、コンヴェンションに対する知識を持つとともに、コンヴェンションについて論じる資格を持つ主体の可能性を見ているとは言えるのではないだろうか。もしこれが正しいならば、制度について論じ、制度に対して批判的に関与したり、制度を再設計していく可能性を制度論に含める可能性が開けるだろう。

*16 制度に関する不可謬主義については、Guala (2016, 第 11 章) を参照されたい。

*17 ここでは、この難しい哲学的な概念区別には立ち入らない。ルイスによれば、この区別はアベラールによるものである。合成的な意味で一般的というのは、たとえば、すべてのドライバーが右側を運転するというときに、その一つの期待を持っていることを意味する。これに対して、分割の意味で一般的というのは、各ドライバーについて、そのドライバーが右側を運転するというような多数の期待を持つことを意味する。

*18 ここで私は通常の古典論理学において論理を最初に定式化する際の言語のことを想定している。

6 コンヴェンションの規範性について

第 III 章はコンヴェンションに類似した諸概念とルイスのコンヴェンション概念との対照を行っている箇所である。類似的な諸概念とは、すでに掲げた目次にあるように、合意、社会契約、規範、ルール、同調的行動^{*19}、模倣の 6 つである。ここでも、すべてを取り上げることができないので、私自身が特に注目する「規範」との対比について、ルイスが述べていることを見ていくことにする。

6.1 「コンヴェンションは規範の一つの種である」という主張

ルイスが目指している主張は明快である。「コンヴェンションは規範の一つの種 (species) であるかもしれない。人が同調すべきだと、われわれが信じるような規則性である」(p.97)。もちろん、彼が与えたコンヴェンションの定義の中に規範性が一切含まれていなかったことを彼は認めている。「わたしがコンヴェンションに与えた定義には、『する義務がある (ought)』『すべき (should)』『善い』、その他の規範的用語が含まれていなかった。また、どの等価な定義を見ても、われわれが規範的用語が生じることを期待する理由は存在しない。したがって、わたしの分析では『コンヴェンション』それ自体は規範的用語ではない」(ibid.)。にもかかわらず、コンヴェンションが規範の一つの種であると主張するのはどうしてなのか。早速、彼の論証を見ていこう。

ある規則性 R が状況 S での行動における、集団 P のコンヴェンションであるとき、以下のことが蓋然的に成立すると、ルイスは言う。

- (1) わたしとともに状況 S に関与する P の他のメンバーのほとんどは、 R に同調するだろう。
- (2) わたしとともに状況 S に関与する P の他のメンバーのほとんどが R に同調するならば、わたしもまた同調することを選択する。
- (3) わたしとともに S に関与する P の他のメンバーのほとんどは、理由を持って、わたしが同調することを期待する。
- (4) わたしとともに S に関与する P の他のメンバーのほとんどは、彼らのほとんどが同調するならば、わたしも同調することを選択する。
- (5) わたしは (1)–(4) が成立していることを信じる理由を持つ。

ここで「蓋然的」と述べているのは、単にその前の第 II4 節の議論において、彼がコンヴェンションの度合いを定義し、コンヴェンションの定義のいたるところで出てくる全称量化子「すべて」を「ほとんどの」に弱めるコンヴェンションの概念化を提案しているからである。したがって、この議論のロジックを理解する局面においては、「蓋然性」という言葉はそれほど意味を持たず、無視しても構わない。むしろ重要なのは、これらがいずれもコンヴェンションの定義の項目の一部になっていることである。

議論の中心はここからである。これらが成立しているならば、以下のことが (蓋然的に) 成立する。

- (6) わたしは、わたしの同調がわたし自身の選好に合致するだろうと信じる理由を持つ。
- (7) わたしは、わたしの同調が、わたしとともに S に関与する P の他のほとんどのメンバーたちの選好に合致するだろうこと、そして、わたしが同調することを彼らが期待する理由を持つことを信じる理由を持つ。

ルイスはこの推論をどのようにして導いているのか。正確なところはわからないが、第一に「コンヴェンションの知識」でなされた議論—コンヴェンションは共通知識として定義されており、しかも共通知識については、

^{*19} この概念はシュウエイダによるもので、まぎらわしいが “comformative behavior” であって、コンヴェンションに同調して行われる “conforming behavior” とは異なる。大雑把に定義すれば、同調的行動とは『それがルールだ』という理由による行為 (p.118) のことをいう。

そのことを全員が期待する理由を持つという議論—が用いられていると考えることである。そして第二に、コンヴェンションの定義の中で一貫して、一様な同調がコーディネーション均衡になっていると考えられていることである(たとえば、本稿の定義3の(3)を見よ)。このことは、上の(1)-(4)には反映されていなかった。(6)と(7)の二つを認めることで、次のような結論が導かれる。

(6)と(7)は、それらが真であるときに、なぜわたしが同調すべきなのかの推定上の(presumptive)理由となる。というのは、他の事情が等しい限り、人は自分自身の選好に合致することをすべきだと、われわれは推定するからである。そして、他の事情が等しい限り、われわれは、人は他の人たちの選好に合致することをすべきであると推定するからである。その人がそうするだろうと、他の人たちが正当に期待するかもしれないときには、とりわけそうである。このとき、どんなコンヴェンションに同調するどんな行為に対しても、われわれは、なぜそれがなされるべきなのかということの、これら二つの(蓋然的かつ推定上の)理由を認識することになるだろう。わたしの理解する限り、われわれは、なぜそれがなされるべきでないのかに関しては、同様に一般的な理由を認識しないだろう。コンヴェンションを規範の一つの種と呼ぶことでわたしが意味しているのは、このことである。(p.98)

ルイスはコンヴェンションに従う行為について、それがなされるべき他の理由が存在するかもしれないことを認めている。たとえば「約束」は約束した人の行為を拘束するかもしれない。この場合には、行為がなされるべき、なされないべきというさまざまな推定上の理由は、互いに秤にかけられるという。その結果、(6)と(7)は簡単に覆されるかもしれないのだが、それでも主体の熟慮には入らなければならないのである。

これまでの議論は、「どんなコンヴェンションも、定義によって、人が同調すべきであるという何らかの推定が存在するような規範である」(p.99)とまとめられるだろう。これに駄目押しするかのようになり、次いで「わたしはここで、それが定義によって、社会的に実効化された(socially enforced)規範でもあることを主張するだろう」と言い、「人は同調することを期待されており、同調しないことは他の人から非好意的な反応を引き起こす傾向にある」という。他人はわたしのことをどう見て、どのように反応するのだろうか。ルイスは続けて、以下のことが蓋然的に成立するという。

(8) わたしとともに S に関与する P のほとんどの他のメンバーたちは、わたしが同調すると期待する。

(9) わたしとともに S に関与する P のほとんどの他のメンバーたちは、条件(1)-(5)が成立すると信じる理由を持つ。

そして、(9)が成立するとき、以下も成立するという。

(10) わたしとともに S に関与する P のほとんどのメンバーたちは、わたしの同調がわたし自身の選好に合致するだろうと信じる理由を持つ。

(11) わたしとともに S に関与する P のほとんどのメンバーたちは、わたしの同調が彼らの選好に合致することと、彼らがわたしの同調を期待する理由を持つことの両方を、わたしが信じる理由を持つことを、信じる理由を持つ。

したがって、わたしの非同調を観察したときに、「彼らはおそらく、わたしが故意にわたし自身の選好に反して、そして故意に彼らの選好と彼らの正当な期待に反して、行為したと推論する立場にあるだろう」(p.99)。こうして非好意的な反応が引き起こされるのである。確かに、コンヴェンションに違反したときに、われわれの内部に引き起こされる日常的な現象学を巧みに説明しているといえよう。

6.2 規範性についての議論を支えているもの

ここでは、ルイスの規範性の議論について、四つのコメントを列挙しておく。

まず、第I章におけるコーディネーション均衡という概念の導入がここにおいて大いに関連性を持っている

ことを指摘しておきたい。通常のゲーム理論に慣れ親しんでいる読者であれば、第I章でコーディネーション均衡の概念が登場する場面で、なぜナッシュ均衡ではいけないのかという疑問を持つに違いない。実際、先に引用したギンタスも述べているように、その後の議論の大半は(厳密)ナッシュ均衡を前提とするだけでも成立するように思われるからである。しかし、私はこの問題への回答が、規範性に関するこの部分の分析において初めて明らかになると考える。

ルイスが「なぜそれがなされるべきなのかということの、二つの(蓋然的かつ推定上の)理由」をあげるとき、そこには、「わたし」が自分の選好に従って行為を選択していることに対する期待だけでなく、主体が属する集団の他のメンバーたちの選好に合致する行為を選択することに対する期待が含まれている。そして、このことが主張できるのは、コンヴェンションの定義において採用されているのがナッシュ均衡ではなくて、コーディネーション均衡だからである。他のメンバーたちが正当な理由を持って期待しているときに、それに反する行為を選択すれば、他のメンバーたちからの非好意的な反応をもたらすと考えられているが、それは単に期待に背いているということだけでなく、彼らの選好に背くことをも意味しているのだ。そして、ここで議論されているように、誰もがそのことを信じる理由を持っているのである。『コンヴェンション』全体を通して見ても、この箇所以外にコーディネーション均衡であることが特に重要な関連性を持つ箇所は見当たらない。

第二に、上記の点とも大いに関係することだが、ここでコンヴェンションに規範性を導入するルイスの議論が、ヒュームが『人間本性論』第三巻「道徳について」で行なっている議論とかなりの程度類比的であり、規範性が自己利益と他者の利益に関する感覚を基盤として説明されていることを指摘しておくべきだろう。ヒュームは第三巻第二部第二節「正義と所有の根源について」の冒頭において、「正義の規則が人間の人為によって確立される仕方」と「その規則を守ることを道徳的美とし、おろそかにすることを道徳的に醜悪であるとするようにわれわれを決定づける理由」に関して議論することを明確にしているが(Hume, 1740/2012, 邦訳 p.39), このうちの第二の問いがここでのルイスの議論に関連している。

ヒュームは正義に伴う「自然な責務(責務を課す性質)」と「道徳的な責務、つまり正と不正の心情」を区別している。まず前者については、利益によって十全に説明されると、ヒュームは述べる。すなわち、「人々がこの諸規則を設定し、一般的にも、すべての個別の事例でもそれに従うように動かされるのは、最初は利益への顧慮のみによってである」(Hume, 1740/2012, 邦訳 p.53)と言う。だが、これだけでは、秩序を維持することの利益を見失いがちなわれわれの傾向性を抑えるには不十分である。われわれには、自分の利益を目の前にするとき、秩序の維持によって得られる利益を見失う傾向があるからである。しかし他方では、われわれは他人の不正義によって被る損害に関しては敏感であって、他人の不正義を観察したときには不快感を感じるものである。この不快感こそ「悪徳」と呼ばれるものなのだ。われわれは社会の中の人々の不快感を共感によって分かち持つことで、道徳的善悪の感覚を持つようになる。こうした他人の正義と不正義に対する感覚は、自分自身の行為にも拡張されるようになる。これが「道徳的な責務」を生み出すわけである。

ここで、規則からの逸脱が自己の利益を損う—そのことは別の手近な利益によってしばしば見失われるのだが—と考えられているとともに、他者の利益をも損うと想定されていることに注目すべきである。ヒュームの議論はしばしばゲーム理論以前にゲーム理論的な推論を行なっているとして“avant la lettre”と形容されるが、まさにその真骨頂が表われている部分である。ルイスの議論には「共感」という概念こそ登場しないものの、基本的にヒュームによる道徳的感情の議論と軌を一にしていることが了解される。

ルイスが「規範の一つの種」と言うとき、おそらく、ここで説明していることが規範性一般の説明になっているとは考えていないだろう。したがって、規範性、道徳性一般が同じような仕方でも説明できると考えてはいなかったかもしれない。しかし少なくともここでの議論が利益ベースのものであることは紛れもない。簡単に言うならば、ルイスはコンヴェンションからの逸脱が自己の利益と他のすべての人々の利益をともに損うものであることに着目し、それが他者からの非好意的反応を引き起こすことに規範性の要素を見出している。

第三に、ルイスは規範性を基本的にゲームの均衡としての性質にかかわらせているものの、その仕方は、他のゲーム理論的な規範性の主張と比較して、それほど単純なものではないことにも注意すべきである。ゲーム理論家の中には、社会規範をゲームの均衡と見なす立場が存在している。たとえば、Osborne and Rubinstein

(1994) は繰り返し囚人のジレンマにおける均衡戦略を社会規範として解釈している。それほど明確に述べていないものの、進化ゲームの理論家たちも進化ゲームの均衡を一種の規範と考えているように思われる。ルイスのここでの議論は、ゲーム理論ベースでありながらも、人々のコンヴェンションにおける知識によって、同調に対する是認と非同調に対する否認が生じるとしているところにその独自性があると考えられる。ルイスの議論はその意味で、ゲーム理論による単純な議論を一段洗練された議論になっているが、逆に、一般化された是認を道徳性の起源と考えるヒュームとの関連で言うならば、ルイスの議論はそのメカニズムをゲーム理論的考察を行うことで、さらに深化させているとすることができるだろう。

第四に、ルイスの規範性の議論に対して、なお残される疑問についても述べておこう。上述したように、規範性をゲームの均衡によって説明する立場にもさまざまに異なるものがあるとはいえ、ルイスの立場が基本的にはこの立場に所属するものと見なされるべきである。そして、規範性をゲームの均衡と見なす立場の多くが、コーディネーション・ゲームに焦点を当てる傾向にあるのだが、この点においてもルイスはまったく同じ傾向性を有していると言うべきである。しかし、規範性の議論をコーディネーション・ゲームの分析だけに基づくものとしていいのだろうか。もう一つの有力な候補として存在しているのが、囚人のジレンマである。囚人のジレンマ的状况に直面するとき、われわれが協力する性向を持っているということは多くの実験研究によって明らかにされてきたことである。ゲーム理論においては、繰り返し囚人のジレンマにおいて協力行動が均衡解として説明できることはよく知られている。しかし、囚人のジレンマの実験結果は、繰り返しではない状況を被験に徹底的に知らせたとしても、人々が協力する傾向にあることを示している。

囚人のジレンマ的状况における協力を規範的と見なすことに反対する人はそれほど多くないと思われる。またこの問題については、夥しい数の研究がなされてきたことも事実である。にもかかわらず、決定打は存在していない。この問題を、制度をゲームの均衡と見なす立場によって説明することはきわめて難しい状況にあると言えるのではないだろうか。この点はルイスだけでなく、ヒュームについても当てはまるだろう^{*20}。

7 結びにかえて

本稿は、これまで頻繁に論じられながらも、その議論のロジックや性格をそれほど詳細に検討されることのなかったデイヴィッド・ルイスのコンヴェンション論について、『コンヴェンション』の前半部に主要な焦点を置きながら検討してきた。冒頭でも述べたように、ルイスのコンヴェンション論はこれまでに多くの制度論者によって取り上げられてきたが、近年のゲーム理論研究の中で、これまでほとんど省みられてこなかったルイス独自の共通知識概念の分析が再評価されるようになっている。本稿において私は、こうした新たな研究潮流の存在をも念頭においたうえで、ルイスの議論が制度論の文脈で果しうる役割を見定めようとしてきた。

ゲーム理論を経由した制度論の現在の視点から、ルイスの議論がどのように見えるのかについては、すでに行論中にコメントを差しさむような仕方で述べてきた。しかし、本稿を閉めるに当たって、それらのうち重要と思われるいくつかの点を整理しておこう。

まず、経験の共有(共通知識の基底)、帰納的推論、背景情報、合理性の期待といった諸概念を駆使した共通知識の概念化の独自性である。ゲーム理論が数学的に定義されたゲームの中に閉じた議論しかしないのに対して、ルイスの議論は、ゲームの均衡をプレーするために必要な整合的な相互的期待を、われわれがどのようにして獲得するのかを説明しようとしている。こうした議論の仕方は、Schelling (1960) にも類比的であるが、制度論から見た場合には、われわれが周囲の環境を作り上げることで、われわれの認知的付加を環境にオフロードすることが得意な動物であるという視点と大きな親和性を持っている (Clark, 1998)。

さらに、コンヴェンションの定義に際して、共通知識概念を適用したことによって、コンヴェンションその

^{*20} ヒュームは囚人のジレンマ的状况を考えていたとも言われており、それを伺わせる箇所は存在する。たとえば『人間本性論』第三巻第二部第五節「約束の果たす責務について」にある例がしばしばエピソードとして引用されている。しかし、コーディネーション・ゲームと囚人のジレンマの違いを明示的に理解していたかどうかは疑わしい。当然ヒュームはゲーム理論を知らなかったわけだが、その限界がここに表現されているとすることができるかもしれない。Heath (2008, 邦訳 pp.94-98) は、この点に関してヒュームを批判している。

ものに対するわれわれの知識について語るができるようにしたことも重要である。そうすることで、制度がどのように維持されるのかを説明できるだけでなく、制度変化にとって知識が果たす役割を分析できる道が開かれることになるかもしれない。しかし、ルイスのコンヴェンションの最初の定義に見られるように、コンヴェンションのコアにあるのは「行動の規則性」であることも見逃してはならない。ルイスのコンヴェンション論において、行動レベルと知識レベルは分かちがたく結びついているのである。行動レベルと知識レベルとの密接な結びつきは、もしかしたら制度改革がどのような形で行われなければならないのかに関して、大きなインプリケーションを持つかもしれない。単に知識レベルで批判することだけでも、行動レベルで変わるだけでも、制度変化は完遂しないだろう。

その他にも、ルイスの議論の独自性のいくつかは本稿のはしばしにコメントしてきた。正当化のロジックと習慣的行動との関係、主体が合理的であることに対する期待、コンヴェンションの知識を決して不可謬的なものと見なさず、経験的に獲得するものと見なす観点等々である。また、彼は自分の議論を一種のモデルのようなものと見なしており、経験的に反証可能であるとしていることも興味深い論点である。

他方では、ルイスのコンヴェンション論を制度一般の理論と見なすことに対しては慎重であるべきである。すべての制度がコーディネーション・ゲームによって包含されていると見なすのでなければ、ルイスの議論が光を当てている制度的現象がわれわれを取り巻く制度的現象の一部でしかないことに注意すべきである。このことは、ゲーム理論的に表現するならば、コーディネーション・ゲームと囚人のジレンマとの違いによって、もっともわかりやすく表現することができるだろう。ルイスはコーディネーション・ゲームに焦点を当てて、コンヴェンションが維持されるメカニズムを明らかにし、その規範性を説くことに成功している。しかし、囚人のジレンマ的状况における規範性の存在については、最初から分析の範囲に入っていないのである。ルイスの議論をより広い制度的現象に適用するには、さらなる一工夫が必要とされるのである。

制度論の観点からルイスのコンヴェンション論を評価すると言った場合には、当然私なりの制度論の視点、あるいは他の制度論者のルイスに対する立ち位置の評価等々がかかわってくることになるだろう。いくつかのコメントには多少とも私の立ち位置が感じられたかもしれないが、本稿ではルイスの議論そのものの紹介に止めておきたい。青木、ギンタス、グアラなどの経済学からの制度論の評価を含めて、私なりの見方を提示することは次の課題としておきたい。

参考文献

- Aoki, M. (2001) *Toward a Comparative Institutional Analysis*, Cambridge, MA.: MIT Press, (邦訳：青木昌彦『比較制度分析に向けて』瀧澤弘和・谷口和弘訳，NTT出版，2001年).
- (2011) “Institutions as Cognitive Media between Strategic Interactions and Individual Beliefs,” *Journal of Economic Behaviour and Organization*, Vol. 79, pp. 20-34.
- Aumann, R. (1976) “Agreeing to Disagree,” *Annals of Statistics*, Vol. 4, pp. 1236-1239.
- (1987) “Correlated Equilibrium as an Expression of Bayesian Rationality,” *Econometrica*, Vol. 55, pp. 1-18.
- Aumann, R. and A. Brandenburger (1995) “Epistemic Conditions for Nash Equilibrium,” *Econometrica*, Vol. 63, pp. 1161-1180.
- Clark, A. (1998) *Being There: Putting Brain, Body, and World Together Again*, Cambridge, MA.: MIT Press, (邦訳：アンディ・クラーク『現れる現在：脳と身体と世界の再統合』，池上高志・森本元太郎監訳，NTT出版，2012年).
- Cubitt, B. P. and R. Sugden (2003) “Common Knowledge, Salience and Convention: A Reconstruction of David Lewis’ Game Theory,” *Economics and Philosophy*, Vol. 19, pp. 175-210.
- Gilbert, M. (1992) *On Social Facts*: Princeton University Press.
- (1996) *Living Together: Rationality, Sociality, and Obligation*: Rowman & Littlefield Publishers.
- Gintis, H. (2009) *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*: Princeton University Press, (邦訳：ハーバート・ギンタス『ゲーム理論による社会科学の統合』成田悠輔・小川一仁・川越敏司・佐々木俊一郎訳，NTT出版，2011年).
- Guala, F. (2016) *Understanding Institutions: The Science and Philosophy of Living Together*, Princeton, NJ.: Princeton University Press, (邦訳：フランチェスコ・グアラ『制度とは何か：社会科学のための制度論』瀧澤弘和・水野孝之訳，慶應義塾大学出版会，2018年).
- Heath, J. (2008) *Following the Rules: Practical Reasoning and Deontic Constraint*, Oxford, UK.: Oxford University Press, (邦訳：ジョセフ・ヒース『ルールに従う』，瀧澤弘和訳，NTT出版，2013年).
- Hume, D. (1740/2012) *A Treatise of Human Nature*, Oxford, UK.: Oxford University Press, (第3巻邦訳：デイヴィッド・ヒューム『人間本性論』第3巻道徳について，伊勢俊彦・石川徹・中釜浩一訳，法政大学出版局，2012年).
- Lewis, D. (1969) *Convention: A Philosophical Study*, Cambridge, MA.: Harvard University Press.
- (1975) “Languages and Language,” in Gunderson, K. ed. *Language, Mind and Knowledge*, Minneapolis: University of Minnesota Press, pp. 3-35.
- Osborne, M. J. and A. Rubinstein (1994) *A Course in Game Theory*, Cambridge, MA.: MIT Press.
- Schelling, T. (1960) *The Strategy of Conflict*: Harvard University Press, (邦訳：トーマス・シェリング『紛争の戦略』，河野勝監訳，勁草書房，2008年).
- Sugden, R. (1986) *The Economics of Rights, Co-operation and Welfare*, London, UK.: Palgrave Macmillan, 2nd edition, (邦訳：ロバート・サグデン『慣習と秩序の経済学』友野典男訳，日本評論社，2008年).
- Vanderschraaf, P. (1998) “Knowledge, Equilibrium and Convention,” *Erkenntnis*, Vol. 49, pp. 337-369.
- Young, P. (1993) “The Evolution of Conventions,” *Econometrica*, Vol. 61, pp. 57-84.
- (1998) *Individual Strategy and Social Structure*, Princeton, NJ.: Princeton University Press.
- 青木昌彦 (2010) 「戦略的相互作用と個人予想を媒介する認知メディアとしての制度」，『新世代政策学研究』，第9号。

- (2014) 『青木昌彦の経済学入門：制度論の地平を拓げる』, ちくま新書.
- 飯田隆 (1995) 『言語哲学大全 III』, 勁草書房.
- 倉田剛 (2019) 『日常世界を哲学する：存在論からのアプローチ』, 光文社新書.
- サグデン, R. (2016) 「規範の創発と再生産における顕著さの役割」, 青木昌彦・岡崎哲二・神取道宏 (編) 『比較制度分析のフロンティア』, NTT 出版, 203-221 頁.
- スカームズ, B. (2016) 「社会契約を自然化する際の諸側面」, 青木昌彦・岡崎哲二・神取道宏 (編) 『比較制度分析のフロンティア』, NTT 出版, 181-201 頁.
- 盛山和夫 (1995) 『制度論の構図』, 創文社.
- 野上志学 (2020) 『デイヴィッド・ルイスの哲学』, 青土社.