

新たなる認識論理の構築に向けての試論

——共有知識 (common knowledge) を手がかりに——

鈴木啓司

認識論から見た古典論理の問題点

認識論理は古典論理を拡張した様相論理の一種であるが、認識という様相を加味した場合、古典論理では当たり前であった論理式が通用しない場合が往々にしてある。代入則不可の問題⁽¹⁾はよく知られているが、ほかにもたとえば、こんな例が考えられる。グループGの全要素a, b, cがPという属性を持っていた場合、古典論理では当然、 $\forall x(Gx \rightarrow Px) = P(a) \wedge P(b) \wedge P(c)$ という式が成り立つが、これに「エージェントiは知っている」という意味の認識のオペレーターKiを付加すると、 $Ki(\forall x(Gx \rightarrow Px)) = Ki(Pa \wedge Pb \wedge Pc)$ は成り立たない。これでは何のことやらわからないという人のために日常語に翻訳した例を紹介すると、三人からなるあるアイドルグループGのメンバー全員が未婚者であった場合、古典論理的には、このことは当然、メンバーaは未婚であり、メンバーbも未婚であり、メンバーcも未婚である、ということと同値である。しかし、ここにある認識主体(人間)がいて、彼がグループGのメンバーa, b, cがそれぞれ未婚であることを知っているも、必ずしもグループGのメンバー全員が未婚であると知っていることにはならない。メンバーa, b, cでグループG全員になることを知らない可能性もあるからだ。そのとき彼には未知のメンバーd, e, f…の存在が否定しきれない。よって、グループGの全員が未婚であるとは知らないわけだ⁽²⁾。

この例が示唆している重要なことは、古典論理における全称記号 \forall (すべての)は集合の内部から要素を数えきることと本質上同値だということである(厳密に言えば、数の概念の定義には、また別の作業を加える必要があるが。だから、「全メンバーを見通す」といった方が正確かもしれない)。確かに、数えきった時点ですべての要素はしかじかであると言うことができよう。しかし、対象が数えきれない無限集合となると、古典論理はとたんに馬脚を現す。 $\forall xP(x)$ と $Px_1 \wedge Px_2 \wedge Px_3 \wedge Px_4 \dots$ は永遠に一致しないことになる。これが、実質的に古典論理を土台とするカントールの集合論が数学に持ち込んだ危機の根底にあった問題である。無限集合は数えきれないところに特性があるのだから、それは本来古典論理の埒外にあるはずのものだ。せいぜい扱えても、完成された無限、実無限ではなく、永遠の生成過程の中途段階の有限にとどまる仮無限までである。それをカントールは多少強引に実無限をその枠内に引き込んだといえよう。ではなぜ、その古典集合論で無限という概念が表現可能であったのか。それは、有限集合に対する古典論理の、要素を数えきることすなわち集合の全体という捉えかたにそもそも欺瞞があったからである。いったい、数えきったという判断はどのようにしてなされるのであろうか。それはあらかじめ集合の全体の数を知っていて初めて可能なことであろう。いくつ入っているかわからない箱の中のりんごを数えるときもそうである。りんごの数はまだ知ら

ないが、りんご一箱という集合の全体はすでに知っている。つまり、数えきる前に全体はすでに見切られているのである。また、そもそも数え始める前に何を数える対象として特定するのか。それはあらかじめ対象が何らかの形で定義されていることを要請しよう。ここで当然の疑問がわく。いったい集合とは、要素だけで成り立っているのか、それとも、それらを囲む概念的境界線を指すのか。この外延性をめぐる問題が、ラッセルのパラドクスといった深刻な事態を招来したのであった。要するに、古典論理は集合の内部から要素を一つ一つ数えあげて全体を定義すると見せて、実ははじめに集合の外部から定義となる大枠を内部を包み込むようにはめていたのである。だからこそ、これを土台とした古典集合論では無限も扱えた（もちろん、外から矮小化した形でだが）。しかし、この大枠の根拠を、有限集合なら要素を一つ一つ確認することで後づけ的に与えることができようが、無限集合となるとそうはいかない。畢竟、その大枠はどこから来たのかという根拠づけの問題に悩むことになる。そして、その由って来る外部を、古典集合論ひいては数学は決して内に取り込むことができないのである。

この集合を外部から見るとは、実質的に神の目である。古典論理が神の目の論理と呼ばれるのもそのためである。それは対象が有限集合にとどまっている間は人間にも使い勝手がよかったが、集合論によって無限集合が導入されると、人知を超えた様相を呈してきた。そのことを認識というまさに人知のレベルから問い直し、古典論理に再考を迫っているのが、実は認識論理なのである。だからわたしは、認識論理が古典論理の拡張形である様相論理の一種として発展してきたことに納得がいかない。様相論理は拡張系と位置づけられるように、古典論理

が起こっている出来事を扱う論理ならば、それを取り巻きさまざまな様相（出来事の必然性や可能性など）を対象にする論理である。そのなかに出来事に対する知識や信念といった認識論理が請け負う分野も含まれていることはうなずける。しかし、要諦は、出来事を知っていることではなく、知っているという出来事なのである。つまり、神の目で見た世界（世界の実態）を人間はどれだけ知っているかではなく、知ることによって人間はいかなる世界を作っているのかである。冒頭に挙げた例も、グループ全体を外部から見ている者と内部から見ている者の知識状態の違いが、古典論理では許される同値関係を阻んだのであった。あの有名なゲティア問題⁽³⁾にしても起因はここにある。エージェント自身の知識の自覚と、外部から見たその規定は、必ずしも一致するとは限らないのである。だが、その外部の目も、最終的に神のごとく世界全体を外から見渡すことができるであろうか。それはいうまでもなく不可能である。人間は神になることはできない。古典論理における全称記号 $\forall x$ （すべての x ）、存在記号 $\exists x$ （ある x が存在する）は、究極の外部者、神の視点を暗黙裡に想定したものであった。しかし、英国の観念論哲学者バークリーの有名な言葉「存在するとは知覚されることである」⁽⁴⁾を引くなら、これらの量子子の左横には本来、人間的エージェントの存在を示すオペレーター K_i が置かれてしかるべきなのである。認識論理はそれゆえ、古典論理が長く隠蔽してきた世界を見る視点の問題を明るみに出し、それを論理に取り込もうとする試みといえる。要するに、進むべき道は拡張ではなく、根本を問い直すことなのである。

様相論理が古典論理を土台にしてその扱う分野を広げた拡張系であったのに対して、古典論

理の公理系そのものに手を加え、いわば違ったものの見方、世界観を提示するのが代替論理⁽⁵⁾と呼ばれている一群である。その代表的なものに、直観主義論理がある。これは、簡単な言い方になるが、古典論理から排中律 ($p \vee \neg p$ 「 p であるか、 p でないかどちらかである」) を抜いたものである。この排中律こそ、実に神の視点を保証するものであった。この公理は、ものごととはすべてシロかクロかはっきり決まっている、ということ述べている。これを冒頭の例に当てはめてみると、三人グループの全員が未婚者かそうでないかは、はっきりと分かる。一人一人確かめてゆけばよいのだ。だが、このグループのメンバーが無限にいるとして、どこまでいっても未婚者だけであつたら、このグループ全員が未婚者であると断言できるであろうか。それはできない。無限を数えきすることは不可能だからだ。しかし、無限を見渡せる神の目から見れば、全員が未婚者かそうでないか(一人でも既婚者がいるか)はどちらかにはっきり決まっている。そもそも全知全能の神に、イエスカノーか答えられない問題などあってはならないであろう。だが、あくまで人間の視点に立てば、このようにシロクロははっきりつけられない問題は山ほどある。先の例でも、われわれに言えることは、確認した限りでは全員未婚者であるということだけだ。こうした人間に直観できる範囲内に論理の枠組みを抑えようというのが、直観主義論理である。神の目の論理に対する人間の目の論理だ。新たな認識論理には、このような古典論理の屋台骨に切り込む姿勢が必要であろう。

古典論理と直観主義論理を比べた場合、シntax (公理系) のみならずセマンティクス (意味論) の面でも、人間の認識の様態をあらわすのに後者のほうが優れているところがあるよ

うに思われる。人知の発展のモデルを考えた場合、大ざっぱに言えば、古典論理のほうは世界をあらかじめ設定し、その中でエージェントが徐々に世界についての知識を増やしていくといったものだが、直観主義論理のほうは、人知が自生する樹のごとく次第に成長していくというイメージである。なぜ後者に利点があるかという、前者では世界の枠組みを決めるという根柢なき天下りの視点がまたぞろ要求されているのに対して、後者では人の知識が一步一步着実に下から上に向かって伸びていくさまが表現されているからである。神から見た固定し完成された世界像ではなく、時間とともに発展していく人知の描く世界像(認識史モデルという)のほうが、認識論的にはよりリアルであることは言うを俟たないであろう。

こうしてみると、認識論理は従来の拡張論理の範疇ではなく、代替論理の枠組みの中に入れて再考すべきものであるように思われる。うへではその証左として、認識論の視点から古典論理と直観主義論理を比較したが、わたしはここで、直観主義論理こそ認識論理の名に値すると主張したいわけではない。実は、人知の論理と称する直観主義論理も、人間の認識様態を正確には表していないことが問題なのである。そのとき論点となるのは、やはり排中律である。確かに排中律は人知に照らしてみると強すぎる。神ならぬ人間はすべてのシロクロを知っているわけではない。しかし、排中律を完全に捨て去るのも、今度は人知を矮小化しすぎる嫌がある。われわれはすべてを知っているわけではないが、知らないことがあることは知っているからである。この無知の知ともいべきものを論理の中で正確に表現することが肝要である。直観主義論理のモデルでは、人間は知っていることだけを知っているという意味になって

しまう。その点では、神の視点を背景にもつ古典論理のほうが、われわれが常に抱いている知識の不完全さの自覚を表現できるといえるかもしれない。内部の無知は外部に立って始めて見えるものだからである。

ここまで神という語を多用してきたが、要は、モデルを外から見ている者である。直観主義論理はこの超越的外部者を消し去り、あくまで人知の内部にとどまろうとしたが、この人知が、古典論理が自然に受け入れられているように本質的に集合を外部からあらかじめ規定し、そのとき多かれ少なかれ神の視点に立つものであれば、やはりこうした傾向をも組み入れて構築されるのが、新たな認識論理の不可欠の条件であろう。問題は、暗黙裡に想定されてしまう外部を消し去るのではなく（それは結局不可能なのだ）、いかにそしてどこまでその外部をシステム内に取り込めるかである。次章では、このウチとソトが微妙に絡み合った認識論の有名なトピックを手がかりに、これからの認識論理はどうあるべきかを探っていこうと思う。

共有知識 (common knowledge)

認識論理が一エージェントのみならず複数エージェント（マルチエージェント）の知識状態をも扱えるよう拡張されたとき、そこには共有知識という概念をいかに形式化するかという問題が、当然のごとく起こってきた。共有知識とは、社会的動物といわれる人間が社会を営むうえで基盤としている協調行為 (coordination) が成立するためになくはならぬ条件として、近年、論理学、社会学、経済学、心理学、哲学、コンピューター科学といったさまざまな分野で注目されてきたテーマである。たとえば、社会の秩序を維持する慣習やルールといったものが

機能するには、人々がいかなる知識状態にあることが求められるであろうか。各人が従うべき行動規範を知っているだけではだめである。それは、この行動規範に従うべきことを皆が知っていることを皆が知っていなければならないのである。具体例でいえば、自動車は左側通行をするというルールが事故なく実践されるには、ドライバーであるわたしがそのルールを知っていて、さらに対向車のドライバーもそのルールを知っていて、そして、そのことをわたしが知っていなければならない。でなければ、わたしは安心して左側通行をしようとは思わないであろう。しかし、話はこれで一件着落とはいかない。今言ったことは当然相手にもいえるので、相手が安心して左側通行を選んでくれるには、私が安心して左側通行を選ぶことを相手を知っていてくれる必要がある。そして、相手がそのことを知っていることを知って改めて私は安心して左側通行を選ぶ。そして、改めて相手が安心して左側通行を選んでくれるために、今言ったことを相手を知っていてくれる必要がある…以下無限に続く。こうした現象は何も慣習やルールにとどまらない。協調を旨とする日常のコミュニケーションの場でも、根底に横たわっていることである。たとえば、友達に対して「あの映画どうだった」と問いを発する状況を考えてみていただきたい。これは「あの映画」がどの映画を指すか友達も知っているとなわしがつ知っていて、さらにこのことを友達が知っていて問いにちゃんと答えてくれるとなわしがつ知っていて、さらにわたしのこの期待を友達が知っていてちゃんと答えてくれるとなわしがつ知っていて、さらに…。このように改めて見直してみると、われわれが社会のなかで日々行っているコミュニケーションというものは、実に奥の深いものだという感慨を覚える。各人が個

別に共通の知識を所有している状態から、『皆がこのことを知っている』ということを知っている」という共有知識にいたることで、無限ともいえる段差を飛び越え、社会の基盤をなす協調行為の成立という劇的变化を遂げるのである。

しかし、ここで疑問が生じる。われわれは実際に協調行動をとるとき、このような無限の推論をしているのであろうか。それはとてもありそうにない。もし、しているのであれば、いつまでたっても共有知識は成立せず、社会はいわばフリーズ状態になるであろう。また、論理学でこれを定義する場合、無限の論理式という有限のステップに収まらない厄介な問題が生じる。マルチエージェントシステムの認識論理では、 Ep (everybody knows p 皆が p を知っている) と Cp (p is common knowledge p は共有知識である) というオペレーターを加えるが、これを使って共有知識を定義すると、 $Cp = p \wedge Ep \wedge EEp \wedge EEEp \wedge \dots$ という無限連言の形になる。認識論理における共有知識の形式化は、まさにここをどう処理するかにかかっている。そしてそれが、新たな認識論理構築の要請にもつながってくるのである。では次に、認識論理では定番の共有知識をめぐるいくつかのパズルを例に、この問題を掘り下げてみよう。

マッディーチルドレン・パズル (muddy children puzzle)

マッディーチルドレン・パズル (文字通りには、泥をつけた子供のパズル) とは、こういうものである。今複数の子供が車座になっているとする。その中の何人かは額に泥をつけているが、本人には分からない。そこに先生がやってきて、「この中の少なくとも一人以上は額に泥をつけています。自分の額に泥がついているこ

と分かる子はいますか。」と問うなら、額に泥をつけている子供が全員イエスと答えるには何回この問いを発しなくてはならないか。これはちょっとした論理パズルで、まず一人だけの場合、その子の視野には誰も泥をつけている者はいないので、即座に、自分こそが泥をつけている当事者だということが分かって手をあげる。二人なら、一度目の問いには答えられない。泥をつけている子供にも泥をつけている他の子供が一人以上見えているから、自分が一人以上の別の一人なのかどうか判断できない。しかし、泥をつけている子供同士は互いに相手が答えられないのを見て、「彼が答えられないのは、彼にも泥をつけた者が見えているからだ。でなければ、彼は自分が唯一泥をつけた者であることがわかって、イエスと答えていたはずだ。そして、僕の目には彼以外に泥をつけている者は見えないので、彼が見ている泥をつけている者は、他ならぬこの僕だ」と判断して、二回目の問いかけには互いにイエスと答えるであろう。以下、三人、四人と増えていっても、前回答えられなかったのはこうした事情でという同様の推理を土台に、三回目の問い、四回目の問いに、泥をつけている子供はイエスと答えられる。すなわち、泥をつけている子供の数を n とするなら、全員がイエスと答えるには、 n 回の問いを発する必要があるというわけだ。もちろん、これはパズルであり、状況は非常に理想化、形式化されている。現実には、そこに居合わせた子供が全員このように論理的な思考ができるとは限らないし、二人以上の場合、自分の額に泥がついていることが分かるのは、 $n-1$ 回目の問いが発せられた後だ。だから、しばらく時間を置けば、互いに相手が答えられない状況を確認して、 n 回目の問いを発しなくとも、おずおずと手をあげるということがあるかもしれない。

しかし、その場合、どれだけ時間を置けば互いに相手が答えられないことを確認できるかという、現実問題としてもはっきりとは決められない。だから、これはそれらもろもろの複雑な実情を捨象した、あくまで思考実験上のパズルなのである。

さて、このパズルはどういうところが、共有知識を考えるうえで示唆に富むのであろうか。考えていただきたいが、子供らは先生の「この中の少なくとも一人以上は額に泥をつけています」というアナウンスメントがなければ、先生の問いが何度繰り返されても答えることができない。先生の宣言があって始めて、泥をつけた子供は先生の問いかけにイエスと答えられるのである。この先生の宣言の果たした役割を考えるために、泥をつけている子供が二人の場合を想定してみよう。いま子供らは全員、額に泥をつけている者を一人以上（泥をつけた本人は一人、その他の子供は二人）目にしている。すなわち、先生の宣言が告げている内容は、すでに皆が知っている。なんら新たな情報を含まない先生の宣言が、どうして状況を一変（答えられなかった子供らが答えられるようになる）させたのであろうか。これは一般に、コンウェイ・パラドクス (Conway paradox) と呼ばれているものだが、その答えが共有知識なのである。すなわち、命題 p 「この中の少なくとも一人以上は額に泥をつけている」が、先生の宣言によって、 E_p から C_p に変わったのである。先の子供らがたどった推論も、相手もこのことを知っている、だからこう行動するはずだ、だから僕はこう行動する、といった「相手の知識の知識」のうえに組み立てられていたことが見て取れよう。それは言い換えれば、相手の思考に対するある種の「信頼」である。まさに、「皆がある事実を知っている」が『皆がある事実を知ってい

る』ことを皆が知っている」に変わることによって協調関係を生み出す共有知識誕生の劇的現場を、このパズルは如実に物語っているといえよう。

このように、共有知識を生じさせるのはその知識の内容ではなく、その伝えられ方である⁽⁶⁾。それはいかなるものか。この問題は、マルチエージェントシステム研究の第一人者、ハルパーンやモーゼスらによって詳細に論じられているが、共有知識成立の最重要の条件は、同時性 (simultaneity) である⁽⁷⁾。同時性とは、情報が同時にグループのメンバー全員に伝わることである。その典型的な例が、先のパズルにあったように、メンバー全員を前にしたアナウンスメントである。この状況下により、メンバー全員は知識が共有されたという、今現在自分たちが置かれている状況そのものを知る。それにより、「相手の知識の知識」が得られ、協調行為を可能にする「信頼」が生まれる。そう、共有知識とは、共有知識が成立しているということを含む知識なのだ。だから、無限連言を避けた論理式ではこう表現される。 $C_p = E(p \wedge C_p)$ これを見ても分かるように、皆は p という事実を知っていると同時にそれが共有知識になっていることを知っている。すなわち、共有知識とはメンバー間にじわじわと広がってやがて達成されるものではなく、一気に同時に達成されるのである。少しでも時間差があれば、あの無限連言が始まり、集団は共有知識に到達することは永遠になく、フリーズしてしまうといえよう。

この同時性を考えるうえで、もうひとつ示唆に富む有名なパズルに触れておこう。一斉攻撃問題 (coordinated attack problem) と呼ばれているものである。今、谷の両側に A, B 二つの軍隊が陣取っているとしよう。彼らは協力して明朝早々、谷底に野営する敵軍を襲撃したい。

だが、手勢からいって攻撃は一斉に行われなければならない。少しでも時間がずれたり、片方が参加しなければ敗北は必至である。そこで、Aの隊長が、明朝六時に攻撃開始の伝令をBの隊長に宛てて送る。しかし、伝令が確実に相手にたどり着く保証はない。途中で敵軍につかまるか、何らかの事故に見舞われる可能性がある。そうなれば、一斉攻撃にはふみきれない。Aは安心できない。そこで、Bの隊長は伝令を受け取ったら、了解の伝令をAに送り返す。だが、この伝令も確実に届くとは限らない。Bは安心できない。そこで、Aは了解の伝令を受け取ったら、了解の了解の伝令をBに送り返す。しかし、この伝令も途中で…といった具合に、永久に両者の間で伝令を送りあっても、明朝六時一斉攻撃の共有知識は成立しないのである。これはネットワークの信頼性にかかわる問題だが、その信頼性のゆらぎも、要は情報伝達のタイムラグに起因する。つまり、そのタイムラグの間に情報が改変されたり、失われたりする危険性は、どんなにネットワークの信頼性を高めても払拭しきれないのだ。このパズルも、最初にどこかで隊長同士が顔つき合わせて打ち合わせ内容を確認しあえばこと足りる。二人同時に共有知識成立の場に身をおけば、共有知識は成立する。無限応答の堂々巡りを回避するのは、最初の一瞬が勝負なのである。

共有知識の成立には同時性が必須であることを見てきたが、しかし、この同時性なるものは本当にあるのだろうか。マッディーチルドレン・パズルをもう一度思い出していただきたい。このパズルは状況を理想化したものであるとして各子供の論理能力の差は捨象したが、実はこの差も同時性に深くかかわっているのである。いったい、同時に同じ場所で同一の情報を受け取ったとしても、その意味するところをまった

く同じに同時に二者が理解するということがありえようか。また、 n 人の泥をつけた子供が自分の額に泥がついていることを理解するのは $n-1$ 回目の問いと n 回目の問いの間であるといったが、それはそのインターヴァルのいつの時点か。いつ彼らは互いに相手がイエスと答えられないと判断して、自分の額に泥がついていることに気づくのか。瞬時にか、一分後にか。それは各人の思考能力によるであろう。こうしてみると、彼らの脳内のニューロンの伝達速度にまで分け入れば、決して同時性は実現していないことが分かる。大げさに聞こえるであろうか。ならば、一斉攻撃問題の二人の隊長にしても、相手の脳内までは読めないとするなら、今自分が認識していることがらを相手も同様に認識していると、どうして確信できよう。相手は何か心配事から気もそぞろで、話をよく聞いていないかもしれない。最近精神に変調をきたして話をも曲解しているかもしれない。また、じつは敵方のスパイで、作戦の裏をかく有力な情報が手に入ったと思っているかもしれない。こうした疑念は日常でも自然であろう。さらに付け加えるなら、物理的な次元からもこの宇宙に同時性はありえない。それを述べているのが相対性理論である。二つの地点（ローカルポイント）で現象が同時に起こると言うには、それを測るグローバルな時計が必要である。しかし、そのような絶対時間を測る時計などこの宇宙には存在しない。各観測者は各座標系の上に立ち、そこでは座標系独自の時間が流れているのである（もちろん、それらはばらばらの別個の時間ではなく、ローレンツ変換という作業を施すことによって統一的に扱えはするが）。あなたとわたしが同時刻に同一の情報を受け取るということは、厳密にいったいではないのだ。

かように、同時性が理論的にありえな

いのであれば、共有知識の理論はどうなるのであろうか。単なる机上の空論か。分かり合っていると思っているわれわれは何か勘違いをしているのであろうか。実際のコミュニケーションの現場ではいったい何が起きているのか。確かに理論と現実にはいくばくかのギャップがあるのはうなずける。そこで、共有知識の理論的条件を少し緩め、その近似値を与える方向が考えられる。たとえば、時間を連続的に考えるから同時性はないのであって、ある幅の最小ユニットをもつ離散的なものと考えれば、その間に収まる事象同士は同時とみなしてよい。あるいは、最初のタイムラグは認めても、始まってしまった無限連言あるいは無限応答を、求められる精度にあわせて適当に止めればよい。100%ではなく99.999999%で満足するのだ。実際、さまざまな応用面ではこれで何も不足はない。大体、われわれが一瞬のうちに無限連言の計算をやっているとはとても思えないではないか。

しかし、こうした近似値も、最小ユニットをどのくらいにするか、計算を何桁で止めるかといった恣意的要素が入ってくる部分は、あくまで理論の厳密性を追う者には不満の残るところである。理論面で無限連言を避ける試みの定番は、数学の不動点定理 (fixed point theorem) を使うものである。不動点とは、境界を含む集合 (閉集合) X があるとして、 X から X への自己連続写像 f の $f(x)=x$ を満たすような点 x を言う。グラフ上でいえば、関数 $f(x)=y$ と $x=y$ の交点だ。比喩的にイメージ化すれば、境界ではっきり囲われた領域があって、その中の部分から部分への割り当て (写像) を考えた場合、ある部分は再帰的に自分を割り当てることに必ずなる。つまり、自分から他の部分に動かないのだ。これは日常の現象にも当てはめられるも

ので、たとえば、かき混ぜられたコーヒーカップの中のコーヒーを思い浮かべてもらいたい。カップの中は閉じた空間である。そして、コーヒーは激しく渦巻いている。だが、渦の中心には静止している部分がある。また、地球という閉空間の中にかかる台風の目もそうであろう。かように、この定理には広い応用性があり、ゲーム理論のナッシュ均衡点、理論経済学のミニ・マックス定理ともつながってくる。そして、共有知識がまさにこの不動点であるというのだ。無限連言を避けた共有知識の論理式を思い出していただきたい。 $Cp=E(p \wedge Cp)$ はまさに再帰的な形をしていた。共有知識は、 $EG(p \wedge x)$ (G はグループをあらわす。不動点が閉集合に存在するものなので) の最大不動点なのであった。それは要するに、こういうことだ。複数のエージェントとそれを取り巻くある状況を集合として、状況の部分集合を各エージェントの知識状態に写像していくと、状況と知識状態がぴったり一致して動かない点がある。それはひとつとは限らないが、それらすべてを含む最大の点が共有知識なのである。共有知識とは、共有知識という状況下にあることを皆知っていることであった。ちなみに、最小の不動点は「虚偽」(false) である。虚偽は状況のどこにも写像できないことを思えば、確かに不動点の一種と考えてよいであろう。

不動点定理により、無限連言による共有知識の定義は回避されたかに見える。だが、実は両者は同値なのである。後者が、 $p \wedge Ep \wedge EEp \wedge EEEp \wedge \dots$ といった具合に、下から上に向かって共有知識を定義せんとしたのに対して、前者は集合内の知識状態の上限を設定し、 $X_0 \supseteq X_1 \supseteq X_2 \supseteq \dots$ といった下向きの包含関係でこれを捉えようとする。いうなれば、無限級数の収束値のようなもので、そこにいたるまでに無限

のステップを踏む必要があろうとも、ゴールさえはっきり見えていれば、そして、そこにいたる手順をしっかり把握していれば、それは到達したも同然なのだ。しかし、そうとはいえ、この定義にしても神の目を持つ古典論理の視点に立っていることは否めない。エージェントを取り巻く状況とは、いったい誰が決めるのであろう。どこからどこまでが集合の範囲内に入るのであろう。マッディーチルドレン・パズルの場合、子供らにとって、みな同じ場所にいること、みな人間であることなども共有知識であるはずだ。これらは考慮に入れる必要はないのであろうか。こうしてみると、状況の集合（世界のあり方）というものが、きわめて超越的な視点から恣意的に切り取られていることが分かる。不動点というかなりはっきりしたイメージを与えられても、それが存在する閉集合をどう境界づけるのが明確にされていない限り、やはり超越論的とのそしりは免れえないのである⁽⁸⁾。

新たな認識論理の構築に向けて

以上述べてきた古典論理モデルによる共有知識像の問題点は、どこにあるのであろうか。それによれば、どうしても本質的に無限連言による堂々巡り（再帰的定義による自己言及も同断であろう）を払拭しきれない。これは、エージェントと世界をそれぞれ独立したものとして分けて考えるところから発しているように思われる。古典論理の知識モデルは、世界がまずあって、その中にエージェントが存在し徐々に世界についての知識を広げていくというものであった。そこにはどうしても、世界全体を見渡しその真理を見極める神の目が想定されている必要がある。そもそも、世界全体や神といった次元にまでいかずとも、エージェントが置かれ

たある状況をモデル化する場合、論理学者なりシステムエンジニアなりは、モデルを外から見下ろす特権的地位を占めているといえる。なるほど、内からは見え外からのみ見える対象を扱う場合は、このモデルは有効である。しかし、ウチとソトが一致した知識の飽和状態ともいべきものを表現するとなると、限界を露呈するのである。その飽和状態が共有知識である。

このウチとソトということで、マルチエージェントシステムの中の代表的で対照的な二つの知識形態、共有知識と分散知識を見てみよう。分散知識とは、ある事実をグループの各メンバーは知らないが、それらメンバーが個別に知っていることを持ち寄ると、その事実が推論できるというものである。すなわち、その事実は分散された形で潜在的にグループ内に知られているといえる。たとえば、太郎は花子が一郎か次郎と恋仲であることを知っていて、三郎は花子が次郎と恋仲ではないことを知っていたとすると、太郎と三郎の知識から、花子が恋仲であるのは一郎であることが分かる。しかし、太郎と三郎はこの事実は知らないのである。かように、分散知識は明らかにグループの外からのみ見え定義できるものであるといえる。次に、分散知識と共有知識の中間段階ともいべき E_p では、 p という事実はグループ内のメンバー全員に知れ渡っている。その意味で、ソトはいくらかウチに取り込まれている。しかし、各メンバーは、この事実をメンバー全員が知っていることは知らないのである。最後に共有知識では、共有知識という状況下にいることをメンバー全員が知っている。すなわち、共有知識というモデルの中にいることをエージェント自身が知っているのである。事態はもはや、モデルの外に立つ特権的視点を許さない。ウチからも

ソトは知られている。あるいは、ウチとソトの眺めは一致している。それを回避してあくまでモデルの外に立とうとするかのごとく古典論理は、特定の状況という閉集合を囲って、その中の不動点という枠に共有知識を押し込めようとするが、そんな恣意的な境界づけを超えて共有知識は（パズルの子供らの共有知識には「互いに人間である」ということも含まれうることを想起）、なにやら人間社会全体を基礎づける知識状態として、モデルを作る論理学者をもウチに取り込んでいるのである。古典論理の枠内で共有知識を定義づけようとする場合、どうしても顔を出すあの無限連言は、つねにソトの視点を設定する古典論理の方法論と、ソトがウチで飽和状態になった共有知識の実態との間の、どこまでいっても調和しない根本的な齟齬を表しているように思われる。

以上のセマンティクス（モデル論）からの議論に加え、シンタクスの面からも同様の批判ができる。古典論理の命題はすべて命題だけで自存し、それを拡張した認識論理においても、認識主体はすでに在る命題を知る形になっている。真なる命題は（「知る」ことができるのは、すべて真なる命題である）、神の視点を支えにして、それだけで独自に存在しうるものなのである。だが、認識主体を離れて独立して在る命題などというものが本来考えられようか。命題とは人の頭に宿って始めて存在色を帯びるものではなかろうか。神以外の誰にも認識されず、無色透明な姿でアイデア界をぶかぶか漂っているかのごとき命題に、われわれが何のかかわりであろう。また、それは同時に認識主体のほうにも言えるのであって、いかなる知識もない空っぽの容器のごときものに主体としての存在価値を認められないことは明白である。こうしたシンタクスが生む、命題という情報があたかも

ウィルスかなにかのごとくエージェント間に伝播していくというイメージが、共有知識の同時性を前にして支障をきたすのは明らかである。ウィルスは一瞬にしては広まらない。必ずいくらかの時間がかかるものである。こうしてみると、古典論理の命題の左横にはすべからず、 K_i もしくは $\neg K_i$ が置かれてしかるべきなのである。

ゆえに、新たなる認識論理のシンタクス構築に当たっては、古典論理のトートロジー（論理的真理）の左に K_i が置けるか置けないかの検討から始めるのがよからう。そこでまず問題になるのが、先にも触れた排中律である。すでに述べたように、排中律は K_i をかぶせるには強すぎる。だから、人知の視点に立つ直観主義論理はこれを除いてしまった（くだいようだが、これは簡便のための言い方である）。しかし、その直観主義もそれにより、今度は人知をあまりに矮小化してしまった。人は「知っていることだけを知っている」のではなく、「知らないことがあることも知っている」のである。それにつけて、直観主義の知識モデルでわたしがかねがね不満に思うのは、知識が時間とともに発展していくその理由となるものが組み込まれていないところだ。ソトを一切設定しない自生する知識の樹は、いかなる力で伸びていくのか。本物の樹にしても、外界から栄養分を吸収して成長しているのである。その点では、古典論理モデルのほうがまだ納得がいく。世界があらかじめ存在するのであれば、その中にいる知的エージェントは徐々に周りの知識を獲得していくであろうからである。知識の木が伸びる推進力となるべきものをウチに組み込む必要がある。それが排中律に代わる、より人知的に表した公理である。そしてその候補となるのが、人間のあくなき知的好奇心を生むきっかけとなっている

「知らないことがあることを知っている」であると、わたしは考える。以後、これを仮に「推進知識」と呼んでおこう。

残念ながら現段階では、まだ具体的な公理系を完成するには至っていない。そこで本稿では、その足がかりとなる試案を述べるにとどめる。今言及した「推進知識」は、いかなる形で論理式化できるであろうか。その内容はより正確にいうと、「何か知らないが知らないことがあることを知っている」である。かように表現にこだわるのも、従来の認識論理にも無知の知と言いうる公理はあるのであって、それとの違いを明確にするためである。いくつかある認識論理の公理に S 4 と S 5 がある。前者は肯定的内省 (positive introspection) と呼ばれているもので、 $Kip \rightarrow KiKip$ (「エージェント i は p を知っているなら、それを知っていることを知っている」) という形をしている。すなわち、知識の自覚性をうたった公理である。これに対し後者は、否定的内省 (negative introspection) と呼ばれ、 $\neg Kip \rightarrow Ki \neg Kip$ (「エージェント i は p を知らなければ、それを知らないということを知っている」) というものである。こちらは少し分りにくいかもしれないが、たとえば、一郎は知人の太郎の姓が山田であることを知らない場合、一郎は太郎の姓を知らないことは知っている (自覚している)。もちろん、太郎の姓が山田であることを知っているのは、一郎以外の第三者、突き詰めれば、この論理式を記している者である。あるいは、 p を真偽の分からない何か数学の未解決問題 (たとえばリーマン予想) としたら、数学者 i は p であること (リーマン予想が真であること) を知らないが、その知らないということは知っている。この場合は、人類の誰も p であるか否かは知らないことになる。いずれにしても、 p は「太郎の姓は山田

である」なり、リーマン予想の内容であるなり、真偽の判定できる (あるいはできるかもしれない) 具体的な命題である。しかし、われわれが問題にしている無知の知は、「何か知らないが知らないことがあることを知っている」である。これが公理 S 5 で表される無知の知とは大きく異なることはお分かりいただけるであろう。すなわち、誰も内容を知らないものを、S 5 のように p として具体化して扱うことなど許されないのである。

人間の認識形態を自然に表す公理としては公理 S 4 までが許容範囲で、S 5 は強すぎるとよく言われる。うえに見た例では至極われわれの知識状態に合致していると思われるが、なぜそう言われるのであろう。それは、この公理がエージェントを超越する視点に p の内容を託しているケースがあるからと考えられる。たとえば、コンピューターシミュレーションの分散システム内の各エージェントは、自分の知識の足りない部分を探索できねばならないが、それは何か欠けている知識があることをプログラマーによって知らされていることを意味する。古典論理の例に漏れず、内部者には見えないが外部者には見える事実が設定されているのだ。人間世界に対する神、シミュレーションに対するシステム構築者の視点を、この公理は背景に含みうる可能性がある。だからこそ、強すぎるといわれても、認識論理は一般に (特に人工知能研究の分野では) S 5 を加えた形で論じられるのである。

しかし、われわれが目指すのは、もはやこうした外の視点を暗黙裡に設定しないシステムである。S 5 とは別に「推進知識」そのものを正確に表現する公理が必要なのだ。だが、「知らないこと」を形式化 (形にする) のは、いうま

でもなく非常に難しいことである。ではどうすればよいのか。そこでわたしが提案するのは、時間概念の導入である。「知らないこと」は、今は具体化できないがやがて具体化されうものと捉えるのだ。実は、時間概念はすでにマルチエージェントシステムには取り入れられている。マッディーナルドレン・パズルを論じた際にみたように、共有知識は時間とのかかわりが非常に強いものであった。そこで当然、さまざまな時間概念を表すオペレーターが必要となってくる。それらは、□（「つねに」）、◇（「いつかは」）、○（「つぎに」）、U（「まで」）といったものだが、これらを使って「推進知識」は表せないであろうか。たとえばこうだ。Ki ($\exists x \diamond KiPx$)「エージェント i はいつか P であることを知る x があることを知っている」。しかし、これでは依然として P が現時点で何かはっきりした属性となってしまう。「推進知識」を表すのに時間論理が及ばないのは、この部分だ。すなわち、やはり様相論理の一種である時間論理は、時間を因果律のもとに見通してしまうため、未来に起こることは現時点で推論可能な範囲にとどまってしまうのである。これも、全体を見渡す外の視点を設定する古典論理モデルのなせる業であろう。古典論理学者は自己の創る小世界において、森羅万象を見通すラプラスの悪魔のごとき位置を占めているのである。

時間概念はむしろ、発展する知識像をモデルに持つ直観主義論理にこそ似つかわしい。実際、それは時間軸上の知識状態の諸段階で表される。存在記号 \exists はここでは、古典論理のように「世界の中にある」のではなく、「知識状態の中にある」という意味になる。そのため、未知なるものが前者のごとく純粹に客観的な具体化を迫られることはない。しかし、かように「知っていること」にもとづく直観主義論理にお

いて「知らないこと」を表現するのは、古典論理と同じく、いやむしろ古典論理以上に困難なことはすでに触れた。未来の知識状態にしても、古典論理の真理概念に代わって「証明可能性」を置いたため、それはより論理的展開の範疇におさまることになった。知識状態によってさまざまに枝分かれする認識史はすべて、アルゴリズム的手法によって見通せるのだ。それは、世界の中を流れる古典論理の時間概念よりも、人知自体の論理的成長過程という意味で、不確定性のない固定した時間の流れといえる。そこでは将来情報が加わることで新たな知識状態が生じるといっても、加わることでそれは初めて知識となるのであり、今現在その未知性は完全に考慮の外である。そこには「命がけの跳躍」(salto mortale マルクス『資本論』)は一切ないのである。

これらを総合すると、現段階ではまったく見えず、次の諸段階ではじめて見える事実を示す表現手法が求められる。しかし、知らないことを形にするという、まったくパラドキシカルな試みは、そもそも可能なのであろうか。この難問は実は、なにげない論理式 $\exists xP(x)$ 「P なる x が存在する」にすでに内包されている。この x とはいったい何なのであろう。それはいわば無色透明なあるものであって、そこに P という属性が貼りついてはじめて存在可能になったのである。すなわち、われわれが見ている世界とは、神の目にのみ映る原存在に、人間に認識可能な部分的属性の色がついて構成されているといえる。世界の実態、モノソノモノといった、古代ギリシャから続く西洋認識論の究極のテーマは、現代論理学において記号により形式化されようとも、所詮、神様に丸投げされているというのが実情である。そして、認識論を突きつめていく過程で浮かび上がったあの「知らな

いこと」とは、古典論理が形式的に処理した(隠蔽した)、モノノモノたる x それ自身なのである。その実態は不可知であろうとも、われわれはいかなる部分的属性も持たぬ全体的 x の概念を知っている。表現すべきは、 P の背後に隠れていたこの x である。それにはどうしたらよいか。今のところわたしには、 Ki ($\exists x \diamond Kiy$) という具合に、 P の述語の部分も変項に変えることぐらいしか思いつかない。前者 x はいわゆる束縛変項であり、「知識状態の中に存在する」という範囲限定がつく。それに対し後者 y は自由変項で、認識されるまでは一切限定がつかない。知って初めてそれと分かるものである。これらの意味するところは、「エージェント i は、いつか何かであることを知る何かが存在することを知っている」となり、これが現段階における「何か知らないが知らないことがあることを知っている」の形式的表現の試案である。これはもちろん、従来の論理解釈ではまったくナンセンスだが、今まで本稿で述べたことに鑑みれば、それなりに意味を成すことが納得できるであろう。要はモデルの問題であって、この論理式を充足するモデルを作ればよいのである(それは確かに難しいことであろうが)。この式は結局、 $x=y$ ということであって、 x (モノ) と y (認識) はイコールだという意味である。 y にあえて述語をあてはめるとしたら、それは \exists 、まさに「存在する」そのことになるのか。これがすべての土台となる。この「推進知識」が既成の認識論の公理とどう関係してくるのかはよく分からないが、公理 $S5$ を弱める作用があるのではないかと考える⁽⁹⁾。また、公理 T 、 $Kip \rightarrow p$ 「 i が p を知っているなら、 p は真理である」(これは真理公理 truth axiom と呼ばれ、知っている内容は真理であることに限られるとする、知識と信念を分ける重要な公理である)

の解釈を、言うまでもなくより主観的内在的なものに変えるであろう。だが、公理的アプローチに関しては、まだまだ道半ばというのが正直なところである。

それでは、そのモデルはどのようなものになるのだろうか。直観主義の認識モデルで足りない部分は、今まで論じてきた「推進知識」と、個別のエージェントの知識状態を視野に入れていない点である。人類の認識史といっても、それは集団的なものではない。もともとコントロールの集合論がもたらした危機に対処すべく数学分野から発したこの論理は、数学知識の発展、具体的には、ある理想の(だが人間的な)数学者の思考の展開をモデルとしている。そのため、認識論理はむしろ、世界に相対して世界を知る個々のエージェントという形で、古典論理の枠内で展開されやすかったといえる。そこでこの点を補うべく、人知全体を各エージェントの知識状態に分けよう。エージェントとは個別の知識状態のことであって、エージェントの外に知識対象である世界が広がっているわけではない。シンタクスのところでも触れたが、独立したエージェントも命題も存在しない。 Ki と p はつねにワンセットで認識の最小ユニットを形作っているのである。各エージェントは「推進知識」によって各自の知識を広げていく。それら小集合の集合が発展する人知である。そこでは当然、さまざまな知識状態が拮抗することになる。その中で人知を代表する知識はいかに決定されるのであろうか。古典論理の「真理」、直観主義論理の「証明可能性」に代わる、その判断基準は何であろう。わたしには、そこに共有知識がかかわってくるように思われる。クーンのパラダイム論以来おなじみの話だが、真なる学説とは、さまざまな経緯があるにせよ、結局その時点でその分野の科学者集団に広く受け入

れられている（共有されている）説のことである。そうしたコミュニティの中で、知識状態の同じエージェントはいわばクローンの関係である。クローン同士にはもはや、エージェント1や2といった個別性は存在しない。ひとつの知識状態である。もちろん、ここでは現実における全人格的な同一状態を言っているのではない。問題となるある局面での知識状態である。その局面が大きくなって、多数のクローンが参画する状態になったとき、人類を代表する人知といったものに近づくのであろう。その際の知識の伝播は、ある知識状態が他の知識状態に出会って、後者に知識を受け渡すのではなくこれを感化するという形で捉えられる。ここで注意すべきだが、人知を個別に分けるといっても、その人知は決して先取的に想定されるべきではない。地道にエージェントを一つ一つ集合させていって、その先に見えてくるものでなくてはならない。また、その中には当の論理学者自身も含まれなければならない。彼は決して世界の外に立つ者ではない。世界とはあくまで、彼を観測点としたモデルなのである。そうすると、さまざまな観測者によるモデルが林立可能となるが、モデルの構成法さえ同じであればよいのである。すなわち、絶対的な一つの世界像を構築するのではなく、各観測者の視点に立ったモデルの統一的な構成法を提示するということなのである。これは各エージェントのアイデンティティーの存続を無視できない協調行為、たとえば分担作業の場合の共有知識を考えるうえでも必要である。エージェント1はAの作業を、エージェント2はBの作業を担当するという共有知識では、それぞれのアイデンティティーの確認は重要である。そのとき、知識状態は同じでも、それがよって立つ観測点の違いがこれを保証するのである。要するにこのモデ

ルは、「同じ知識状態の者が出会い、そのことを互いに確認しあう」という共有知識の形を、できうる限り自然に表現しようというものである。「出会い」と捉えれば、それは二つのもの間の現象である。両者の間にタイムラグはない。

最後に、このモデルから見た共有知識の実例はどう解釈されるか、触れておこう。マッディーチルドレン・パズルの子供らは、先生の宣言により、全員が『この中の一人以上は額に泥をつけている』ことを全員が知っている」ことを知った。その知識状態を視覚的にあらわせば次の表ようになるであろう。メンバーはaからeの五人とする。一番左の列が、「当のメンバーは知っている」という意味で、行はその知っている内容である。pはもちろん、「この中の一人以上は額に泥をつけている」という命題である。

Ka	Kap	Kbp	Kcp	Kdp	Kep
Kb	Kap	Kbp	Kcp	Kdp	Kep
Kc	Kap	Kbp	Kcp	Kdp	Kep
Kd	Kap	Kbp	Kcp	Kdp	Kep
Ke	Kap	Kbp	Kcp	Kdp	Kep

かように全員同じ知識状態にあるが、この表自体が共有知識をあらわしている訳ではない。この表があらわしているのは、各自が相手もこの事実を知っていることを知っている状態で、各自のこの知識自体が相手に知られていることは表現されていない。それを表に入れ込もうとすると、またぞろ、あの知識オペレーターの無限連鎖に陥ることになる。共有知識が成立しているのは、この表を一同が一緒に見ているその場である。そしてその視点は、今この表を見ている読者の視点と重なる。繰り返しになるが、共

有知識の形式化の難しさは、このソトの視点をいかにウチに取り込むかということにかかっているのである。新たな認識モデルは、先述したように、同じ知識状態のエージェントはクローンと考える。そのクローン同士が顔を合わせると、エージェント a, b といった区別はなくなり、同じ知識状態のもとに統合される。a や b といった区別があるから、自分がこのことを知り、相手もこのことを知り、そしてそのことを自分が知り、また、このことを相手が知り、といった、共観の無限キャッチボールが始まるのである。新たなモデルは、エージェントと知識状態を同一視することでこれを回避する。その視点からいうと、パズルの子供らの中の額に泥をつけている者が先生の問いかけに答えられるのは、彼らだけは他の子らと、見えている泥をつけた子供の数が違うからである。他の子が泥をつけた子をすべて見ているのに対して、彼らの見ている数は常に一人少ない。彼らの知識状態はその点で、共有知識成立の場においてさらに互いに同調しているといえる。その同調状態が先生の問いかけによって、推論という合理的過程を経るにせよ、彼らの共有知識となって現れるのである。

以上、古典論理を土台とした従来の認識論の限界点と、直観主義論理の視点からのその修正、そしてその更なる改良形の試案を述べてきた。果たして、いっさいソトの視点に支えられないシステムは構築可能であろうか。われわれはいつも有限なる絵を見て物事を分かったつもりになっているが、その絵を絵と認識するためには背景となる無限のキャンパスが必要なことを心に留めておくべきである。そのソトたる無限も含めて認識なのである。それが数量的に実際に無限であるか否かということは問題ではな

い。境界が引けない（引ければそれはたちまち絵になってしまう）ということが肝心なのである。故にわれわれは、困難であろうとも、このソトをウチに取り込んだ認識のシステムを構築する必要がある。それはまた換言すれば、ルールを知らずにそれにしたがって動くエージェントのシステムではなく、構築者と同じくらいルールを知って動くエージェントのシステムである。ルールを知っているからこそ、それをときに意識的に変え目覚しい飛躍をとげる、それが人知の大いなる特徴である。それなくしては、外的環境が変化しない限り自らも変わらない動物となんら異なるころはなかろう。こうしたシステムを具体化する作業はまだこれからだが、いずれ近いうちに、幾分なりと公理系やモデルの形式を示せばよいと考えている。

註

- (1) 古典論理では、 $a=b$ で a が F なら b も F だが、認識の次元ではこれが必ずしも成立するとは限らない。たとえば、よく引かれる例だが、明けの明星と宵の明星は同じ金星だが、明けの明星が金星であることを知っているからといって、即、宵の明星も金星であると知っていることにはならない。二つが同じ星であることをまだ知らなかった古代人を思い浮かべてもらいたい。これはフレーゲ・パズルとも呼ばれ、フレーゲ自身は、彼の有名な意味（直接の指示対象、今の場合は金星）と意義（表示されるものに与えられる様態、明けの明星と宵の明星）理論の形で、この問題を処理した。
- (2) これはバーカン式 (Barcan formula) と呼ばれるものに関係してくる問題で、それによれば、 $\forall x_i \dots \forall x_n. Kip \rightarrow Ki \forall x_i \dots \forall x_n. p$ となる。注目すべきは、量子子と認識オペレーターの位置関係で、左式は、すべての要素 x についてエージェント i は p であることを知っているという意味で、エー

ジェントの外側に立ったいわば客観的事実を述べているのに対し、右式は、エージェント i はすべての x が p であるを知っているという意味で、エージェントの内面的な知識状態を表している。要するにこの式は、客観的に見て知っているといえるなら、内面的にも知っているとみなしてよいというものだが、これは背景となる領域が共通の知識である場合には成立するが、知識状態が違えば成立しないことが分かっている。古典論理とは畢竟、神という共通の知識のもとに眺められた世界の道理のことであろう。ちなみに、上の式の左式と右式を入れ替えた逆パーカン式は、エージェントの知識状態にかかわらず成立する。これは直感的にも納得できよう。しかし、対象が無限集合となるとどうか。われわれは、すべての自然数はしじかであるを知っているからといって、一つ一つの自然数すべてがそうだを知っているといえるか。これはこれでまた新たな問題である。

- (3) 伝統的な知識概念を揺さぶるゲティア問題は、1963年、E. L. Gettier が *Analysis* 23号に発表した数ページの短い論文「正当化された真なる信念は知識だろうか」に発する。この認識論史上有名な論文において著者は、従来受け入れられてきた、ある特定の命題を誰かが知っていることの必要十分条件、すなわち、 P だということを S が知っているのは、1) P が真であり、2) P だということを S が信じており、かつ、3) P だということを信じる点で S が正当化されている場合である、に疑念を挟んでいる。そこに示されている二つの反例はいずれも、たまたま真である命題を正当なる根拠を持って信じている場合で、要するに、知識成立の条件には、当人の内面的要素だけでなく、外的状況も加味しなければならないことを示唆している。ここに状況全体を見渡す外の視点が要求されていることは明らかであるが、問題は、その外の視点が最終的に世界全体のソトに立ちうるかということである。それが不可能であるなら(実際、不可能だと思うが)、その視点は誰かのウチのものであろう。ちなみに、この論文の翻訳は、森際康友編『知識という環境』(名古屋大学出版局)、1996、に、柴田正良訳で掲載されている。
- (4) ジョージ・パークリ『人知原理論』大槻春彦訳(岩波文庫)、p. 45。アイルランドの哲学者パークリ(1685-1753)は、知覚を経ない事物の存在を否定した極端な観念論者として知られている。
- (5) 代替論理として他に代表的なものには、古典論理の真偽二値に代わり、その中間段階を認める三値論理、多値論理がある。これも、排中律の拒否から発した論理である。
- (6) メッセージの内容よりその伝えられ方に着目することで、公共的な式典や集会などの「公共的儀式」が持つ意味に新たな光を当てた社会学的成果としては、数年前に翻訳が出た、マイケル・S・Y・チュウ『儀式は何の役に立つのか』安田 雪訳(新曜社)、2003、などが挙げられる。社会的統合と既存の権威体系を維持するのに役立つ公共的な式典、政治的、社会的変革をもたらす公共的な決起集会やデモ、これら協調行為を促す「公共的儀式」は、単に情報の中心から個別の聴衆へのメッセージの伝達ではない。それは同時に、聴衆に他の聴衆が何を知っているかを知らせるのである。(同書 p. 2)
- (7) R. Fagin, J. Y. Halpern, Y. Moses, M. Y. Vardi. *Reasoning About Knowledge* (MIT Press), 1995.
- (8) 不動点定理で救いきれない共有知識の奥深さについて、ゲーデルの不完全性定理とのアナロジーで一言触れておこう。ゲーデルはその定理の証明に当たって、ゲーデル式というものを構成した。これは「あるゲーデル数(全式に構成的な方法でそれぞれつけられた番号のようなもの)を持つ式の証明式のゲーデル数となる自然数は存在しない」といった解釈のできる式のことであり、この式そのもののゲーデル数がとりもなおさず、「あるゲーデル数」となっている。すなわち、この式は「自分の証明式のゲーデル数となる自然数は存在しない」と言っていることになり、すなわち、それは「自分は証明不可能である」というに等しい。彼はこうして、真と解釈すれば証明不可能な式を構成することにより、自然数論を展開できる形式体系の不完全性を証明したのであった。それはそれとして、ここで強調したいのは、あるゲーデル数を代入したその式のゲーデル数が代入した

ゲーデル数そのものになるというゲーデル式の、不動点を表す式 $f(x) = x$ との自己言及的な類似性である。やはり認識レベルの問題にまで拡大解釈された不完全性定理のことを思うと、不動点定理の普遍性もここにかがいがい知れるような気がする。しかし、ゲーデル式が不動点と決定的に違うのは、後者が閉区間に存在するのに対して、前者は自然数列という無限の中に存在していることである。そしてそれにより、自分がどこにも写像できないこと（自分の証明式の番号となる自然数が自然数列の中に入らないこと）が証明できないことである。無限を対象とした直接的な不在証明は不可能である。すなわち、ゲーデル式は不動点らしきものだが、不動点であると証明できない。人間が数を理解する動物であるなら、数学の根本知識はある意味究極の共有知識のひとつであると言えるかもしれない。そうした意味で共有知識は、深く追求すればするほど、ある閉じた領域にとどまるものではなく、広く人間の認識の枠組み自体に繋がるものであるように思えてくる。

- (9) 形式的なことはまたの機会に譲ろうと思うが、様相論理のセマンティクス、クリプキモデルでは、S5はユニヴァーサルモデルに当たる。クリプキモデルとは、別名可能世界論とも呼ばれ、さまざまな可能世界の集合間の到達可能性関係で、必然性や可能性といった様相を表現するモデルである。その関係には推移性や反射性、対称性などがある

が、ユニヴァーサルモデル（別名ユークリッド的モデル）はこれら三つを兼ね備えた一番強力なもので、そこでは、任意の集合から任意の集合へ自由に行き来できる（これより弱いモデルでは、到達可能性に制限がある）。すなわち、ラブラースの悪魔のごとき古今東西すべてお見通しといった視点を獲得できる古典論理の世界である。これに対し「推進知識」は、現時点では見えない（到達できない）未知のものを要請することで、S5に弱い認識論的解釈（「命題化できるはっきりしたことについての無知を知っている」）を与えるものであると思われる。ちなみに、ゲーデル・マッキンゼイ・タルスキの定理により、直観主義命題論理は様相命題論理 S4 に変換埋め込み可能であることが知られている。やはり認識の形式化の鍵は、古典論理と直観主義論理の間、S5とS4の間にあると思われる。

主要参考文献

- R. Fagin, J. Y. Halpern, Y. Moses, M. Y. Vardi. *Reasoning About Knowledge* (MIT Press), 1995.
 戸田山和久『論理学をつくる』(名古屋大学出版会), 2000.
 野矢茂樹『論理学』(東京大学出版会), 1994.
 小野寛晰『情報科学における論理』(日本評論社), 1994.