不完全性定理をよりどころにした人工知能の発展可能性に関する考察

高　橋　信　行

目　　次

論説

不完全性定理をよりどころにした人工知能の発展可能性に関する考察

高橋信行

もくじ

梗概

人工知能の飛躍的発展が報告されている。これらの報告の中には正しいと思われるものもあるが、中にはセンセーショナルに人々をあおるだけの記事もある。そこで本稿では、まず黎明期の人工知能を振り返る。次に、コンピュータによってできることと出来ないことの明確な境目の存在を確認する。更に、このような境目の概念が何ゆえ誕生したのかという問題について考察し、非可解性の概念を生み出した不完全性定理の証明を検討する。更に Turing 機械の停止性判定問題が非可解であることを確認する。以上の考察から、本稿の主題である「与えられた手続きが正しいアルゴリズムであるか否かを判定する人工知能は存在し得ない」という主張を導く。この主張から 2045 年問題で提起されている危惧に対して、ある種の指針を提言する。

キーワード：不完全性定理, 帰納的非可解性, Turing 計算可能性, Church's Thesis, 2045 年問題, 技術的特異点

§0.　はじめに

“人工知能（Artificial Intelligence）”なる用語がはじめて登場するのは、1956 年に米国で開かれたダートマス会議の開催案内においてである。しかし用語の誕生の前に概念の誕生があって、これは 1947 年に Alan Mathison Turing によって提出されている。Turing はその後 1950 年の論文 ”*Computing Machinery and Intelligence* ” の中で図 1 に示すチューリング・テストを提出し“人工知能”の概念を明確にしている。

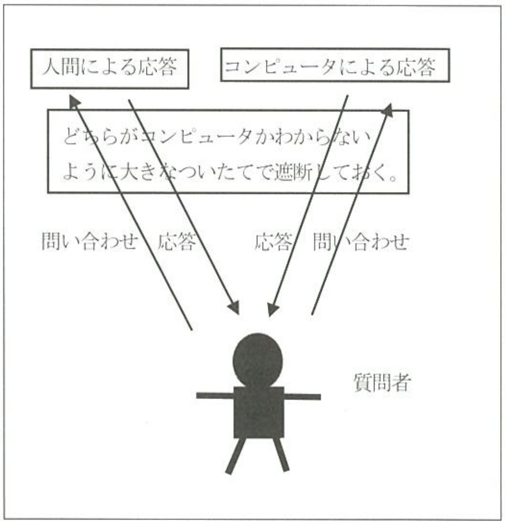

図 1　チューリング・テスト

実際に種々の手続きの中でも高度な知能による処理を思わせるものもあれば、そうではなくて単純な機械的処理と思えるものもある。たとえば成田を立ってロンドンまで飛ぶにあたり、最も安くて最も安全で、しかも最も快適なプランを 1 つだけ提供してくれるシステムがあったとする。利用者は、このシステムに何度も問い合わせをすることが許されているとする。システムが「お客さまには候補となるプランが $p_1, p_2, \ldots, p_m$ だけあります。この中から p_i が最適なプランです。」と提示したとする。確かにこの p_i が正しい解であったとしよう。$p_1, p_2, \ldots, p_m$ の中から最も安いプランを選び出せという問題で、このシステムが p_j を解として与えた場合、単純な最小値を求めるプロ

グラムを走らせれば瞬時に最適解 p_j が与えられる。このように機械によって最適解が得られるような単純なアルゴリズムは、Turingの定義では人工知能とは呼ばない。しかし最も安くて最も安全で、しかも最も快適なプランを p_i として提示したシステムが果たして、本当に機械だけによるものなのか、それとも人間が隠れて答えたのかは定かではない。実は、このシステムは 100%機械によるものなのだが、人間による仕業なのか、それとも機械による提示なのかが判断がつかないときにTuringはこのシステムを“人工知能”と呼ぶことを提唱したのである [1)]。

ここでTuringの定義では、機械によって最適解が得られるような単純なアルゴリズムを使っている場合は、人工知能とは呼ばれないことに注目する。人工知能（Artificial Intelligence）に対して、本稿では数理知能（Mathematical Intelligence）を考える [2)]。両者ともアルゴリズムにしたがって、ある種の仕事を行う点では共通である。人工知能は、人間が工学的に作り上げた言わば feasible procedure（物理的に実行可能な手続き）を遂行するのであるから、時代によってその能力は大きく異なって来る。一般的には、その能力の広がりは単調非減少である。それに対して数理知能は、帰納的手続きのことである。Church's Thesisは広く認められているし妥当な提唱であるから、数理知能とは計算可能な手続きの総称と言ってよい [3)]。この概念は、時代を超えた概念であってすでに 1930 年代に確立された数学上の概念、正確に言えば帰納的関数論におけるKey Stoneである。人間が人工的にまたは工学的に作り上げた手続きは、コンピュータの能力に左右されてしまう訳である。一方、数理知能の全体はアルゴリズムの全体と同一である。将来、誕生するかも知れない量子コンピュータでは実行可能でも、今日のコンピュータでは実行不可能なアルゴリズムも存在する。そこで本稿では人工知能の概念をつぎのように定める。

【本稿での人工知能の定義】人工知能とは人間が工学的に作り上げた知能で、現実に作動しているコンピュータによって実行可能な手続き（feasible procedure）を遂行する仕組みをいう。

この定義にしたがうと先ほどの機械によって最適解が得られるような単純なアルゴリズムも人工知能と呼ばれることになる。Turingの定義では、あたかも人間の行為のような高等な動作を以て人工知能と定める訳であるが、“あたかも人間の行為のような高等な動作”の範囲を定式化することは非常に難しい。

本稿では、人工知能と呼ぶに際し“あたかも人間の行為のような高等な動作”であるかどうかは基準にしない。そうではなくて、工学的に実行可能な手続き（feasible procedure）を遂行する仕組みのことと定める。このように人工知能を定めたとき、人工知能の全体と帰納的に可解な問題群との関係は図 2 のように表される。

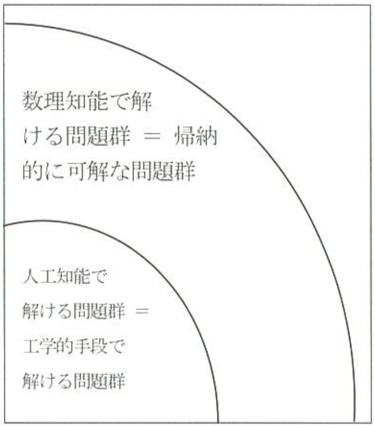

図 2　人工知能で解ける問題群
⊆　数理知能で解ける問題群

ダートマス会議に参加した John McCarthy や Marvin Minsky らの仕事は、この節では触れない。Minsky と Papert の共同作業は、その後の人工知能研究において大きな礎石を残したのであるが、本稿ではこの時代を遡ること更に 20 年、すなわち 1930 年代の Turing と Gödel の仕事に依拠して論を進める。なにゆえ 1930 年代に依拠するのか、といえばこの時代にプログラム内蔵方式のヒントやアルゴリズムの数学的定式化が実現されたからである。コンピュータに関わる技術の源は、1930 年代に求められると言って過言ではない。しかしながらアルゴリズムの定式化という本稿の主張の琴線ともいえる主題が、突然 1930 年代に誕生したのではない。結論からいえば、確かに 1936 年に Church's Thesis として計算可能性についての一応の統一的見解が得られた。しかしなにゆえ、1936 年に Church's Thesis が得られたのか？という問いに対しては、振り返れば 1900 年に開かれた国際数学者会議における David Hilbert による 23 個の未解決問題の提示から論を起こすのが妥当と思える。

§1.　動機と経緯

筆者は横浜商科大学紀要第 4 巻に『SM-計算可能性

と SM-計算可能関数のクラス』と題して、順序機械による計算可能性について報告した。その稿のなかでは、なにゆえ“計算可能性”が議論の対象となるのか、という問題について1920年代に起こったHilbertらによる有限主義の立場から論を起こしている。その論考の中では、当然ながらGödelの不完全性定理についても言及している。ただ、なにぶんにも当時の筆者の考察が不充分で今に至るも十分、満足に足る考察が行われたという意識を筆者自身、持っていない。

ところで今日、人工知能の技術的発展が驚異的である。事実、米Google社は2016年1月27日、同社の開発した人工知能システム "AlphaGo" が囲碁の試合で人間に勝利したことを発表している。Google社ではこの人工知能技術を、気候モデリングや複雑な病疫分析など、現実世界での重要な問題解決に適用していくということである。しかし人々の関心のなかには人工知能が究極的にどこまで進化するのか？そして人工知能はどこまで人間に取って代わるのか？といった漠然とした不安感がある。このような不安をあおるように、人間は人工知能に職を奪われやがて失業者が増大し、不幸な社会がやって来るという記事を目にすることがある。一方、対局的に人間は人工知能にはできない分野で更なる新境地を開き、よりクリエイティブで豊かな生活が実現されるといった肯定的な見方もある。そのような時代背景にあって人工知能の技術が、政治・経済・文化・産業・教育、その他人々の生活の深層に与えつつある影響を無視することはできない。このような一種の地殻変動に対して先人の足跡を押さえ、さらに将来に向けた盤石な視座を確立しておくことは万人にとって意味のあることである。

本稿では厳密な科学的検証、特に§3 における不完全性定理の証明を経て、最終的に下記の A. M. Turing による Entscheidungsproblem 判定不能定理を確認する。

【Turingの定理】プログラム P と入力 x が任意に与えられたとする。入力データ x に対して P の計算が正常に停止するか、それとも無限に続くのかを判定するアルゴリズムは存在しない。

アルゴリズムが存在しないということは、何人もそのプログラムが作れないということである。どの人工知能もプログラムを内在している。つまり、プログラムが作れないということは人工知能も作れない、ということを意味する。また、プログラムが停止しない場合、正しいプログラムとは言えない。したがって、つぎの本稿の主題を得る。

【本稿の主題】与えられた手続きが正しいアルゴリズムであるか否かを判定する人工知能は存在しない。

どんなロボットも人工知能を搭載しているので、この主題からつぎの主張を得る。

【本稿の主張】正しく作動する人工知能を自律的に作るロボットは、存在し得ない。

つまり極めて優秀なロボットが居て次々と優秀なロボットを自動的に生産するような夢のような時代は、永遠にやって来ないのである。次々と人工知能を作り出したとしても、これらの人工知能を作動させるプログラムたちが、停止するのかそれとも停止しないで永遠に動き続けるのかを判定するプログラムは存在しないのである。もしもこの判定作業を実行するとすれば、人間が検証システムを設計する以外に方法はない。

このような主張が生まれる背景には、1936年に発表された Turing[1]がある。§4 において、この Turing の定理の証明を吟味する。そのために§2 で Turing の定理と Entscheidungsproblem（ 決定問題 ）が生まれた時代的背景を探ることにする。

§2.　歴史的経緯

なぜ歴史を振り返るのか、と言われれば、何人も時間軸を一つしか持つことが出来ないからである。SFの世界の話のように過去に戻り、人生をやり直すことは出来ない。しかし近い将来を予測して、より正しいと思われる方向に自分や社会を向かわせる方が、無駄な反省をしないで済む。天気予報などは、その典型である。今日、天気予報の精度はかなり高い。それは、過去の膨大なデータの蓄積とスーパー・コンピュータによる分析と予測技術の賜物である。つまり近未来をよりよく生きるために、過去を振り返ることは極めて有意義なのである。

“はじめに”で述べた通り、Gödelによって得られた不完全定理の証明方法とその意味は、20世紀後半の科学思潮、特に論理学や計算機科学に大きな影響を与えた。ここで 1931 年に発表された所謂、不完全定理を述べておくことにする。

【第1不完全性定理】自然数論を内在する公理系 S が ω-無矛盾であれば、S の中に肯定もその否定も証明できない閉論理式 A が存在する。

後述するが、定理の証明で示されるゲーデル文 A は内容的な解釈のもとで真である。しかし、この体系で証明され得ないのである。この第1不完全性定理から、つぎの第2不完全性定理を導くことができる。

【第２不完全性定理】自然数論を内在する公理系 S が無矛盾であれば、公理系 S の無矛盾性は S の内部で証明できない。

これらの主張は、特に第２不完全性定理は Hilbert のもくろみを完全に打ち下してしまったのである。Hilbert は、1900 年にパリで開かれた国際数学者会議において 23 個からなる未解決問題を提示した。その精神は形式主義の勝利を万全にし、以て Cantor らによる逆理からの超克を実現するためである。1874 年 Georg Cantor は集合論を創始し濃度や超限順序数の概念を作り出した。Cantor の生み出した集合論は、今日公理的集合論と区別するために素朴集合論と呼ばれている。Cantor と Dedekind による素朴集合論は、概念の自由な創生と優れた記述性によって、従来の数学をも再構成し更に新しい知見を生み出す豊饒性を蓄えていた。しかしながら集合論の中に逆理（パラドックス）が、見つかってしまったのである[4]。19 世紀末から 20 世紀初頭にかけて種々の逆理が発見されている。ここでは幾つかの逆理の中でも最も素朴で、しかも本質的と思われる Russell の逆理を取り上げる。

2-1.　Russell の逆理

Bertrand Russell（1872-1970）について、多くを語るには筆者の力量を超える。それほどまでに多才な人物である。通説では Russell から Frege あてに宛てた 1902 年 6 月 16 日付けの手紙が Russell の逆理の起源とされている。しかし、同じ頃 Zermelo もこの逆理を発見している。

【Russell の逆理(1903 年)】まず、認識できる対象の全体 Ω を考える。集合は Ω の要素の中でも、特に共通の特性(property)を持つ対象たちだけで構成される。すなわち、集合はつぎのように定義される。

$$A=\{x\in\Omega \mid P(x)\}$$

ここに $P(x)$ は x の持つ特性(property)をあらわす。どの要素も Ω の要素であるから、いちいち $x\in\Omega$ と記す必要はない。そこで集合 A はつぎのように書かれる。

$$A=\{x \mid P(x)\}$$

つまり P という特性を持つ要素 x の全体を A であらわすわけである。したがって

(1)　$x\in A \quad \Leftrightarrow \quad P(x)$

x の持つ特性 P としては、何でもよい。そこで P としてつぎの特性を考える。

(2)　$P(x) \quad \Leftrightarrow \quad x\notin x$

つまり

$$R=\{x \mid x\notin x\}$$

なる集合 R を考える。R は自分自身を要素に持たない対象の集まりである。R も認識できる対象であるから $R\in\Omega$ である。したがって $R\notin R$ であるか、さもなければ $R\in R$ のどちらか一方が成り立たなければならない。今、$R\notin R$ が成り立つと仮定すると(2) から $P(R)$ である。(1) から $R\in R$ である。一方、$R\in R$ を仮定すると (1) から $P(R)$ を得る。(2) から $R\notin R$ である。以上の議論から

$$R\in R \quad \Leftrightarrow \quad R\notin R$$

を得る。これが Russell の逆理である。このような逆理を避けるべく Russell は型の理論（Type Theory）を創設した。型の理論によって、どの程度の数学を実現させせることができるのかを実証するために Russell は Whitehead と協力して Principia Mathematica：（数学原理）を 1910 年から 1913 年にかけて出版した[5]。Russell と Whitehead のスタンスは、論理主義と呼ばれる。論理主義に触発されて多くの数学者が、公理主義の立場で上記のような逆理を生成しない方策を提案した。中でも Hilbert による形式主義と Brouwer による直観主義との論争は有名である。Hilbert の主張は、数学理論を形式化してその体系で矛盾が生じないことを有限の立場で証明するという所謂、有限の立場である。つまり、証明とは有限個の公理から出発して有限個の推論規則を正しく有限回、適用して得られるプロセスと考える。重要なのは、公理と推論規則をどのように設計するかである。上述の「$R\notin R$ であるか、さもなければ $R\in R$ のどちらか一方が成り立たなければならない。」という主張は、排中律に準拠している。排中律は、いかなる主張 A も not A であるか、それとも A であるかのどちらか一方が必ず成り立つという公理である。この公理を認めないと証明過程の厳密性を優先させる余り、得られる結果の貧困を招く。事実、Brouwer の直観主義では、排中律を認めないので、人類が古代ギリシャ以来 2000 年の間に蓄えて来た多くの定理が証明されない事態を引き起こす。しかし、このような立場での証明を計算過程とみなすことで、プログラムの正当性をチェックするツールを作成できるのである。その意味では Brouwer の見識は、実際のコンピュータの誕生とソフトウェア・クライシスの時代まで埋もれたままであった。1920 年代における Hilbert と Brouwer の論争を経て、多くの数学者は Hilbert の形式主義を支持し、古典論理が論理学における一般的礎石となった。1900 年に Hilbert によって提示された 23 個の未解決問題（unsolved problems）の

中で肯定的に解決されたものは多い。中には、問題自身の曖昧さから未解決のままのものもある。しかし10番目の問題は、今日では非可解問題（unsolvable problems）として決着するに至った。非可解問題とは、その問題を解くアルゴリズムがないという問題群である。そもそも、この“アルゴリズム”という用語はいかなる概念を表しているのだろうか？

2-2. Church’s Thesis

ある表現を借りれば、アルゴリズムとは「有限個の指令の集まりであって、それに忠実にしたがう限り、だれでも創造的な思考なしに有限時間内に所要の結果を得ることができるような手続きのこと[6]」と言える。しかし、この引用文「・・・」の中には指令であるとか、所要の結果、手続きなどおよそ数学的概念とは言えない直観的表現が含まれている。したがって「・・・」のような表現をいくら工夫してみても数学的議論としては成立し得ない。

もちろん、アルゴリズムのある問題に対しては「・・・」で示したような解き方を具体的に示せばよい。しかし、アルゴリズムの無い問題に対し非可解性を示すには、その問題がアルゴリズムの無い問題のクラスに属していることを示さなくてはならない。したがって、アルゴリズムの無い問題のクラスを集合概念としてきちんと定式化する必要が生じるわけである。これは必然的に“アルゴリズム”を上記の「・・・」のような直観的概念としてではなく、集合論的記述のできる数学的概念としてとらえる必要性をもたらす。しかし、“アルゴリズム”が上記のような、あいまいな直観的概念である以上、何か別の数学的概念で置き換えるしか手がない。そこで1936年 A. Church は、アルゴリズムのある関数として帰納的関数を充当させたのである。その結果、その問題を解くアルゴリズムが存在しそうもないときには、その問題が帰納的でないことを言えばよいことになり、これは数学的議論として成り立つ話である。

しかし、

【Church の提唱】:“アルゴリズム”のある関数とは、帰納的関数のことである。

という提唱は、大胆過ぎはしないだろうか？

否、充分根拠のある極めて妥当な提唱なのである。実際に、人類がこれまでに知っている“アルゴリズム”のある関数は、すべて Turing 計算可能でありしたがって帰納的である。また、Turing 機械のテープを k 本に増やしたり、非決定性にしたりしても帰納的関数のクラスは不変である。更には、多くの研究者による種々の“アルゴリズムの定式化”が、すべてこの帰納性に一致したことも Church の提唱の妥当性を裏付けている。たとえば、Emil L. Post は1936年に双正規性の概念を発表し、Turing 機械とほとんど同様な計算モデルを提案している。双正規性が Turing 機械と同じ計算能力を持つことは、自明であった。一方、A. Church は λ-定義可能性を発表し Kleene と Church によって、この概念もやはり帰納性と一致することが証明された。Turing 自身も λ-定義可能性と彼の Turing 機械による計算可能性とが等しい概念であることを証明している。このような事実から“アルゴリズム”のある関数という、我々が直観的にしか捉えられない概念を帰納的関数、すなわち Turing-Post 計算可能関数で代替させようというのが、Church の提唱の意味である。言い換えれば、“アルゴリズムのある関数”の数学的定義と言ってもよい。したがって Church の提唱は、アルゴリズムの定式化という先人たちの研究努力に対する明確で、しかも安定した一つの解答である、ということが出来るわけである。

2-3. 計算不能関数の存在

ω から ω への関数 の全体は、連続の濃度を持っている。

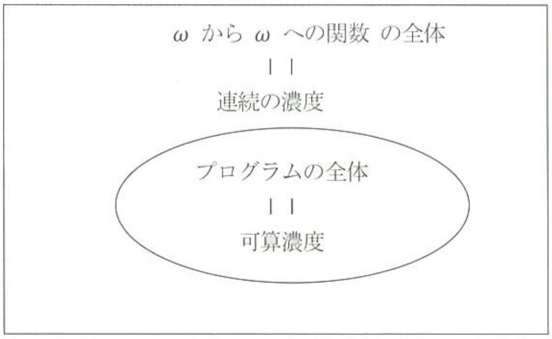

図3　アルゴリズムのない数論的関数の存在

一方、アルゴリズムの全体つまりプログラムの全体は、プログラム1, プログラム2, プログラム3, ... という具合に番号をつけることができるので図3に示す通り可算である。後で述べるが対角線論法を使って、連続の濃度は可算濃度よりも真に大きいことを示す。このように濃度の比較から、アルゴリズムを持たない ω から ω への関数（数論的関数）の存在が言える。図3からはアルゴリズムのある数論的関数よりも、む

しろアルゴリズムのない数論的関数の方がたくさん存在することがわかる。このように非可解性の存在は、Gödel の不完全性定理を持ち出さなくても言えるのであるが、それは Church の提唱をこの議論で使っているからである。前述の Church の提唱がもしも無ければ、つまり帰納的関数の全体とプログラムの全体とが同じ集合であることが認められなければ、図 3 を論拠にする議論は認められない。

やはりトップダウンで議論するよりもボトムアップ方式の方が確実である。そこで§3 と§4 では§2 でたどった回帰を逆に時間軸に沿った形でたどることにする。

§3. 不完全性定理

3-1. 対角線論法

本節では Cantor がはじめて用いた対角線論法を吟味する。不完全性定理も Turing の定理も対角線論法が重要な役割を果たすからである。

自然数の全体 ω から $2=\{0,1\}$ への関数の全体

$$\mathcal{R}=\{\ f\ |\ f:\ \omega \to 2\ \}$$

を考える。$f\ \in\ \mathcal{R}$ は、$f(0)=0, f(1)=1, f(2)=1, f(3)=0, \ldots$のような形をしている。0, 1 の無限列が循環すれば、つまり 0, 1 の有限列が無限に繰り返す場合は有理数とみなせるし、01 列が循環しない無限列の場合は無理数とみなせる。つまり、$\mathcal{R}$ は 2 進数で表された実数の全体 (the set of all reals) である。ただし、次の Cantor の定理における $\mathcal{R}$ は、10 進数で表された実数の全体（the set of all reals）とする。

【Cantor の定理】 ω の濃度、$\mathrm{card}(\omega)$ を可算濃度という。 $\mathcal{R}$ の濃度は、可算濃度ではない。

[証明] 実数の全体 $\mathcal{R}$ が、可算濃度を持つと仮定して矛盾を導く。01 区間 $I=\{x\ \in\ \mathcal{R}\,|\ 0<x<1\}$ と $\mathcal{R}$ との間には、全単射写像 $f(x)=\frac{2x-1}{x(1-x)}$ が存在するので $\mathrm{card}(I)=\mathrm{card}(\mathcal{R})$ である。$\mathrm{card}(\mathcal{R})=\mathrm{card}(\omega)$ と仮定しているので、$\mathrm{card}(I)=\mathrm{card}(\omega)$ となる。したがって 01 区間 $I=\{x\ \in\ \mathcal{R}\,|\ 0<x<1\}$ の要素たちは、つぎの図 4 のように $r_0, r_1, r_2, r_3, r_4,$ ・・・という具合に縦一列に並べられる。ただし $r_i=0\,.\,x_0x_1\cdots x_n$ 0 0 0・・・のように、あるところから先すべてが 0 となる場合は $r_i=0.x_0x_1\cdots x_n'$ 999・・・($x_n'=x_n-1$) として、あるところから先すべて 9 が続く表現の方を採用して表現の一意性を担保する。

$$
\begin{aligned}
r_0 &= 0\,.\,x_{00}\,x_{01}\,x_{02}\,x_{03}\,x_{04}\ldots\ldots\ldots\\
r_1 &= 0\,.\,x_{10}\,x_{11}\,x_{12}\,x_{13}\,x_{14}\ldots\ldots\ldots\\
r_2 &= 0\,.\,x_{20}\,x_{21}\,x_{22}\,x_{23}\,x_{24}\ldots\ldots\ldots\\
r_3 &= 0\,.\,x_{30}\,x_{31}\,x_{32}\,x_{33}\,x_{34}\ldots\ldots\ldots\\
r_4 &= 0\,.\,x_{40}\,x_{41}\,x_{42}\,x_{43}\,x_{44}\ldots\ldots\ldots\\
&\ \ \vdots \qquad\qquad \vdots
\end{aligned}
$$

図 4 $I=\{x\ \in\ \mathcal{R}\,|\ 0<x<1\}$ の 10 進小数展開

いま、新たにつぎのようにして実数

$$s\ =0.\ s_0\,s_1\,s_2\,s_3\ldots\ldots$$

をつくる。

すべての自然数 $i=0,1,2,3,\ldots$ に対して

$x_{ii}\in\{\ 0,1,2,3,4\ \}$ のときは、 $s_i\ =7$

$x_{ii}\in\{\ 5,6,7,8,9\ \}$ のときは、 $s_i\ =2$

とおく。このように定義された s について $0<s<1$ であるから $s\in I$ である。s は I の要素であるから、図 4 のどこかに登場する筈である。つまり、自然数 k が存在して $r_k\ =s$ となる。r_k の 10 進展開の k 番目は、x_{kk}である。ところが s の 10 進展開の k 番目は、s_k であって $s_k\neq x_{kk}$ である。したがって $r_k\ \neq\ s$ となってしまう。つまり自然数 k が存在して $r_k\ =\ s$ とはならない。換言すれば、どんな自然数 k に対しても $r_k\ \neq\ s$ である。このことは s が図 4 の列 $r_0, r_1, r_2, r_3,$ …… の中に決して出現しないことを意味する。したがって$s\notin I$ である。これは不合理である。よって実数の全体 $\mathcal{R}$ が、可算濃度を持つことはない。

【証明おわり】

このように対角線論法は議論の主人公について図 4 の対角線上で定義し、この主人公が舞台からはみ出してしまうようにして矛盾を導く論法である。この論法を使って以下、第 1 不完全性定理を証明する。ただし Gödel 自身は、第 1 とか第 2 という言い方はしていない。“不完全”という表現は後世になってつけられた通称である。第 1 不完全性定理を発表したときの正式な論文タイトルは、“Über formal unentscheidbare Sätze der *Principia matehmatica* und verwandter Systeme I”である。日本語訳すれば「プリンキピア マテマティカ及び類似の諸体系における形式的に決定不能な命題について I」となる。この“プリンキピア マテマティカ”とは、前述の Russell and Whitehead による「数学原理」のことである。「プリンキピア マテマティカ及び類似の諸体系」として、Gödel 自身 *ZF* を挙

げている。このあたりの様子を知るためには、1931年に発表されたGödelの論文を精査することが有効である。以下、原文の冒頭部分を引用し日本語訳を付ける。

" Über formal unentscheidbare Sätze der *Princi pia matehmatica* und verwandter Systeme I[1] "
(1931)

Die Entwicklung der Mathematik in der Richtu ng zu größerer Exaktheit hat bekanntlich dazu gef ührt, daß weite Gebiete von ihr formalisiert wurd en, in der Art, daβ das Beweisen nach einigen we nigen mechanischen Regeln vollzogen werden kann. Die umfassendsten derzeit aufgestellten form alen Systeme sind das System der *Principia mat hematica* (*PM*)[2] einerseits, das Zermelo — Fraenk elsche (von J. von Neumann weiter ausgebildete) Axiomensystem der Mengenlehre[3] andererseits. Di ese beiden Systeme sind so weit, daß alle heute in der Mathematik angewendeten Beweismethoden in ihnen formalisiert, d. h. auf einige wenige Axiome und Schlußregeln zurückgeführt sind. Es liegt dah er die Vermutung nahe, daß diese Axiome und Sc hlußregeln dazu ausreichen, *alle* mathematischen Fragen, die sich in den betreffenden Systemen übe rhaupt formal ausdrücken lassen, auch zu entschei den. Im folgenden wird gezeigt, daß dies nicht der Fall ist, sondern daß es in den beiden angeführten Systemen sogar relativ einfache Probleme aus der Theorie der gewöhnlichen ganzen Zahlen gibt,[4] die sich aus den Axiomen nicht entscheiden lassen. Di eser Umstand liegt nicht etwa an der speziellen Na tur der aufgestellten Systeme, sondern gilt für eine sehr weite Klasse formaler Systeme, zu denen insb esondere alle gehören, die aus den beiden angeführ ten durch Hinzufügung endlich vieler Axiome ents tehen,[5] vorausgesetzt, daß durch die hinzugefügten Axiome keine falschen Sätze von der in Fußnote 4 angegebenen Art beweisbar werden.

1. Vgl. die als *1930b* erschienene Zusammenfassung der Resultate dieser Arbeit.
2. *Whitehead und Russell 1925.* Zu den Axiomen des Systems *PM* rechnen wir insbesondere auch: Das Unendlichkeitsaxiom (in der Form: es gibt genau abzählbar viele Individuen), das Reduzibilitäts- und das Auswahlaxiom (für alle Typen).
3. Vgl, *Fraenkel 1927, von Neumann 1925, 1928a, 1929.* Wir bemerken, daß man zu den in der angeführten Literatur gegebenen mengentheoretischen Axiomen noch die Axiome und Schlußregeln des Logikkalküls hinzufügen muß, um die Formalisierung zu vollenden. ——— Die nachfolgenden Überlegungen gelten auch für die in den letzten Jahren von D. Hilbert und seinen Mitarbeitern aufgestellten formalen Systeme (soweit diese bisher vorliegen). Vgl. *Hilbert 1922, 1923, 1928, Bernays 1923, von Neumann 1927, Ackermann 1924.*
4. D. h. genauer, es gibt unentscheidbare Sätze, in denen außer den logischen Konstanten: $\neg$ (nicht), $\vee$ (oder), $\forall x$ (für alle x), = (identisch mit) keine anderen Begriffe vorkommen als + (Addition), ・(Multiplikation), beide bezogen auf natürliche Zahlen, wobei auch die Präfixe $\forall x$ sich nur auf natürliche Zahlen beziehen dürfen.

以下は、筆者による日本語訳である。訳出に当たっては、前原[7]を参照した。

「プリンキピア　マテマティカ及び類似の諸体系における形式的に決定不能な命題について I 1)」
(1931年)

1

より厳密な正確さを求める数学の発展は、よく知られているように広範な分野が形式化されるという結果をもたらした。そこでは証明は、いくつかの機械的規則を用いて実行される。目下のところ最も包括的に作られた形式的体系は一方ではプリンキピア　マテマティカ *PM* [2] の体系であり、他方では Zermelo-Fraenkel による集合論の公理系 *ZF*（これは J. von Neumann によって、更に発展させられた）である [3]。この２つの公理系においては、今日の数学において使われているすべての証明法が形式化されてしまう程に、言い換えれば少数の公理と推論規則だけに還元されてしまうという程度に *PM* と *ZF* は広範囲なのである。このことからすぐに、ここで議論の対象となっている体系の中でおおむね形式的に表現されたすべての数学上の問題が成り立つのか、成り立たないのかを決定できるほどに、これらの公理と推論規則は十分に具備されているとの予想が成り立つ。以下、本文の中で実はこの予想は事実ではなく、それどころか引き合いに出した２つの公理系の中にはごく普通の自然数上の比較的簡単な問題で、公理系で決定できないものが存在することが示される [4]。この主張は、既に確立されている体系の何か特別な性質に原因があるのではなくて、非常に広い形式的体系(Formal Systems)のクラスに当てはまる事態なのである。上述の２つの体系から有限個の公理を追加することによって生ずる公理を使って、脚注の [4] で挙げられた種類の偽な命題は決して証明され得ないという仮定のもとで、つまり算術を含む体系が矛盾していなければ、とりわけ２つの引用された体系、すなわち *PM* と *ZF* に有限個の公理を追加して得られたすべての体系について、この主張はなりたつのである。

1) この仕事の結果をまとめた要旨が *193cb* として出版されるので、この要旨も参照されたい。

2) *A. Whitehead, B. Russell, Principia Mathematica, 1925.* この体系 *PM* の公理の中に以下の３つの公理を含ませるものとする。無限公理（ちょうど可算個の個体が存在するという形での無限公理)、還元公理、それからすべてのタイプの選択公理の３つを *PM* に追加しておく。

3) *Fraenkel 1927. von Neumann 1925. 1928a. 1929* を参照されたい。定式化を完了させるためにはここに挙げた *Fraenkel, Neumann* による集合論の公理に、論理計算のための公理と推論規則を追加しなければならない。*D. Hilbert* と彼の同僚によって最近、構築された形式的体系に対しても、この論文での主張は成り立つ。*Hilbert 1922. 1923. 1928. Bernays1923. von Neumann 1927. Ackermann 1924* を参照されたい。

4) もう少し精確にいうと以下の通りである。論理記号 $\bar{\ }$ (否定 not), $\vee$(または or), $\forall x$(すべての x に対して for all x), および $=$(等号 equal) に加えて、自然数の上の $+$(足し算)と $\cdot$(掛け算)以外は出現しないような肯定も否定も証明できない命題が１つ以上、存在する。もちろん $\forall x$（すべての x に対して、for all　x）もまた自然数 x に対してのみ、適用されるものとする。

このように何か特別な事情を仮定するわけでもなく、単に ペアノの公理系のような自然数の上の算術を実行し得る公理系と ω 無矛盾性を仮定したとき、真ではあるけれどもその公理系からは証明できない命題を実際に Gödel は作ったのである[1]。その証明法は、証明の仕組みを自然数上の算術の中に埋め込むやり方で、今日では Gödelization （ゲーデルの算術化）と呼ばれている。Gödelization を実現するキーワードとして次節で組み立てるゲーデル数が重要である [7]。

3-2.　ゲーデル数

本稿で用いる形式的体系は、古典論理にペアノの公理系を加えたものとし、この体系を *PA* であらわす。

[定義 1]　閉論理式の集合 K から論理式 A に至る証明（英：proof, 独：Beweis）とは、論理式の有限列 $A_1, A_2, \cdots, A_{n-1}, A_n = A$ のことである。ただし、各 i $(1 \leq i \leq n)$ に対しつぎの(1)から(5)のいずれか１つが成り立つ。

(1) A_i $(1 \leq i \leq n)$ は、K の要素である。

(2) A_i $(1 \leq i \leq n)$ は、*PA* の公理である。

(3) A_j $(1 \leq j \leq n)$ が $A_k \to A_i$ $(k < i)$ という形をしていて、A_i $(1 \leq j < i \leq n)$ は A_k と A_j から推論規則によって導かれた論理式である。

(4) A_{i-1} $(1 \leq i \leq n)$ が $A(a) \to C$ という形をしていて、A_i が $\exists x A(x) \to C$ という形をしていると

[1] 後に Rosser は、ω 無矛盾性を無矛盾性で置き換え得ることを 1936 年に示した。

き、A_i は A_{i-1} から推論規則によって導かれた論理式である。

(5) A_{i-1} $(1 \leq i \leq n)$ が $C \to A(a)$ という形をしていて、A_i が $C \to \forall x A(x)$ という形をしているとき、A_i は A_{i-1} から推論規則によって導かれた論理式である。

ただし(4)と(5)において a は C に出現しない自由変数とする。

このように証明とは、証明可能性をみたす論理式の有限列のことである。最後の$A_n = A$ を定理（theorem）と呼び $\vdash A$ と書く。$\vdash A$ は *PA* で論理式 A が証明可能(provable）であることを表している。論理式は記号からなる有限列である。そこで論理式を構成する記号のうち定数記号の１つ１つに対して、まず奇数を割り当てる。

(i)　記号のゲーデル数

次の表５の対応により、*PA* の１つの定数記号に対して１つの奇数を割り当てる。

表５　定数記号のゲーデル数

PA の記号	0	,	∈	¬	→	∀	(	)
ゲーデル数	1	3	5	7	9	11	13	15

変数記号に対しては、17 以上の奇数を対応させる。このようにして、定数記号と変数記号にゲーデル数を対応させる。注意すべきは、いまのところ偶数は割り当てられていないことである。このことから、任意に自然数 x をとって x が 15 以下の奇数のときは表５の定数記号が復元され、17 以上の奇数であった場合は変数記号が復元されることになる。ゲーデル数が偶数になる場合をこれから定める。

(ii)　論理式のゲーデル数

PA の定数記号または変数記号 $X_1, X_2, \cdots, X_m$ から構成された論理式 $S = X_1\ X_2\ \cdots\ X_m$を考える。各 X_i のゲーデル数が n_i（奇数）であるとき S のゲーデル数 $G(S)$ を以下で定める。

$$G(S) = p_1^{n_1} \cdot p_2^{n_2} \cdot \cdots \cdot p_m^{n_m}$$

ここに p_i は i 番目の素数である。つまり $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $\cdots$ である。

つぎに証明のゲーデル数を定義する。定義１で定めたとおり、証明は論理式の有限列である。

(iii)　証明のゲーデル数

PA の証明 $Beweis = A_1, A_2, \cdots, A_{\nu-1}, A_\nu$ が与えられたとする。ただし(ii)で与えられた各論理式 A_i のゲーデル数は n_i であったとする。$(1 \leq i \leq \nu)$

このとき $G(Beweis) = p_1^{n_1} \cdot p_2^{n_2} \cdot \cdots \cdot p_\nu^{n_\nu}$ として $Beweis$ のゲーデル数 $G(Beweis)$ を定める。

(ii) と (iii) は、素数の冪の有限積という形は同じである。実は (ii) で論理式のゲーデル数を定めるとき、指数はすべて奇数であった。しかし (iii) で証明のゲーデル数を定めるときの指数は、すべて偶数である。より明確に述べれば、(i) (ii) (iii) の違いは以下のようにまとめられる。

(i)　記号に対しては、奇数が対応している。

(ii)　論理式に対しては、$2^n \cdot m$ $(m, n$ は奇数)の形の自然数が対応している。

(iii)　証明に対しては、$2^n \cdot m$ (n は偶数, m は奇数)の形の自然数が対応している。

このことから任意に自然数 x が与えられたとき、x が奇数であれば (i) のケースである。x が偶数のとき、素因数分解の一意性から $x = 2^n \cdot m$ $(n > 0$, m は奇数) なる分解がただ１通りに決まる。もしも n が奇数であれば、(ii) のケースであり、n が偶数であれば、(iii) のケースとして復号 (decoding) も一意的に決められる。

[定義 2]　定数記号 0 と関数記号 s について、項（term）をつぎのように帰納的に定める。

Basis　0 は項である。

Induction Step　t が項のとき、s(t) は項である。

このようにして項 $0, s(0), s(s(0)), \ldots$ を得る[2]。項 $0, s(0), s(s(0)), \ldots$ を特に数項（numeral）と呼び

$$\bar{0}, \bar{1}, \bar{2}, \ldots$$

という具合に自然数に bar を付けて表す。0, 1, 2, … は通常の自然数であるが、数項 $\bar{0}, \bar{1}, \bar{2}, \ldots$ は有限の長さを持つ記号列である。

このようにわれわれの議論の中には、内容に関する議論と形式に関する議論の２つが登場する。これらは、それぞれメタ言語、形式言語で表現される。以下、両者を区別するための表記上の約束を置く。

[2] 帰納的定義は厳密にいうと「Extremal Clause　以上によって与えられるもののみが項である。」との極条項を入れる。

まず形式的体系においては、論理記号

$\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\forall$, $\exists$

を使う。一方、内容を表すためのメタ記号としては

$\bar{\ }$, &, *or*, $\Rightarrow$, $\Leftrightarrow$, $(\forall)$, $(\exists)$

を用いる。更に、$x \leq y$ なる自然数についての関係は $\boldsymbol{R}(x,y)$ のように斜体の太字で表す。一方論理式として x ≤ y を見た場合、すなわち $\exists z\ (x+z=y)$ なる表現として見た場合は、R(x , y) として立体の細字で表す。

【表現定理】任意の帰納的述語 $\boldsymbol{R}(x,y)$ に対して、次の条件をみたす Σ_1-論理式 φ(x , y)が存在する。

(1) $\boldsymbol{R}(n,m) \quad \Rightarrow \quad \vdash \ \varphi(\bar{n}, \bar{m})$

(2) $\overline{\boldsymbol{R}(n,m)} \quad \Rightarrow \quad \vdash \ \neg\varphi(\bar{n}, \bar{m})$

証明については、例えば 田中・鹿島・角田・菊池[3] pp. 76-80 に詳しい記述がある。

[定義 3] 自然数 n は、PA の下でのある証明のゲーデル数になっていて、その証明の最後の論理式、すなわち定理のゲーデル数が m になっているとき、この関係を $\boldsymbol{Beweis}(n,m)$ と書く。

したがって自然数 n が、論理式 A の証明のゲーデル数であるという主張は

$$\boldsymbol{Beweis}(n, G(\mathrm{A}))$$

で表される。この $\boldsymbol{Beweis}$ は帰納的である。したがって表現定理によって Σ_1-論理式 Beweis(x , y) が存在して、任意の自然数 n , m に対して

(1) $\boldsymbol{Beweis}(n,m) \quad \Rightarrow \quad \vdash \quad \mathrm{Beweis}(\bar{n}, \bar{m})$

(2) $\overline{\boldsymbol{Beweis}(n,m)} \quad \Rightarrow \quad \vdash \ \neg\mathrm{Beweis}(\bar{n}, \bar{m})$

が成り立つ。

更に A が証明されるという事実は、A に至る証明が存在する、およびそのときに限られるので

$$\exists y \ \mathrm{Beweis}(y, \overline{G(\mathrm{A})})$$

と表される。そこで論理式 $\exists y\,\mathrm{Beweis}(y,x)$ を Bew(x) と表し "x ist beweisbar."と読むことにする。Bew(x) も Σ_1-論理式である。意味は "x は証明可能である。"ということである。したがって

$$\vDash \mathrm{Bew}(\overline{G(\mathrm{A})}) \quad \Leftrightarrow \quad \vdash \mathrm{A}$$

となる。ここに $\vDash \eta$ は、自然数論の標準モデルで文(閉論理式) η が正しいことを表している。

[定義 4] K を論理式の集合とする。

$$K \vdash \mathrm{A} \quad かつ \quad K \vdash \neg\mathrm{A}$$

をみたす論理式 A が存在するとき、K は矛盾する(inconsistent)という。K が矛盾しないとき、すなわち任意の論理式 A についてもしも $K \vdash \mathrm{A}$ であれば $K \vdash \neg\mathrm{A}$ とはならないとき、K は無矛盾である(consistent)という。

[定義 5] 論理式の集合 K が ω-矛盾する(ω-inconsistent)とは、

$$(\forall n)\ [\ K \vdash \mathrm{F}(\bar{n})\ ] \quad かつ \quad K \vdash \exists x\ \neg\mathrm{F}(x)$$

となる論理式 F(x) が存在することをいう。ω-矛盾しないことをω-無矛盾(ω- consistent)であるという。論理式の集合 K が矛盾すれば K はω-矛盾するので、K がω-無矛盾であればかならず無矛盾である。つまり、ω-無矛盾という性質は無矛盾よりも強い性質である。

3-3. 対角化定理

Cantor の定理で述べた対角線上での議論をここで行う[3]。対角線上では、自分自身 ξ の性質を ξ を用いて言い及ぶことになる。自分自身について述べる文のことを自己言及文(self-reference sentence)という。自己言及文は、論理学では古代からよく研究されて来たテーマである。

【定理 6 (diagonalization theorem)】 F(x) は x 以外の自由変数を含まない論理式とする。このとき

$$\vdash (\ \xi \quad \leftrightarrow \quad \mathrm{F}(\overline{G(\xi)})\ )$$

をみたす閉論理式 ξ が存在する。

[証明] x 以外に自由変数を持たない、すべての論理式をゲーデル数の小さい順に並べたとき、

$$\mathrm{A}_0(x),\ \mathrm{A}_1(x),\ \mathrm{A}_2(x),\ \ \dots$$

と並んだとする。自然数の全体 ω から ω への関数 $\boldsymbol{f}$ を

$$\boldsymbol{f}(n) \ = \ G(\mathrm{A}_n(\bar{n}))$$

で定める。

$\boldsymbol{f}$ は帰納的であるから表現定理によって、$\boldsymbol{f}$ は表現可能になる。すなわち、任意の n に対して

(1) $\vdash \ \forall y\ [\ y = \overline{\boldsymbol{f}(n)} \quad \leftrightarrow \quad \chi(\bar{n}, y)\ ]$

であるような論理式 χ(x , y) が存在する[4]。ただし、χ(x , y) は x , y 以外の自由変数を持たない。また y は x とは異なる変数で F の中には出現しない。ここで論理式

$$\exists y\ [\ \chi(x, y) \ \wedge \ \mathrm{F}(y)\ ]$$

を考える。この論理式は自由変数として x だけを持っているので

$$\mathrm{A}_k(x) = \exists y\ [\ \chi(x,y) \ \wedge \ \mathrm{F}(y)\ ]$$

となるような自然数 k が定まる。いま、この定数 k

[3] 3-3 節と 3-4 節の執筆に当たっては、田中・鹿島・角田・菊池[3]を参照している。

[4] $\boldsymbol{f}$ が 帰納的であるから表現定理によって、$\boldsymbol{f}$ は表現可能になる。$\boldsymbol{f}$ を 表現するための論理式として χ(x , y) を用意した。

を用いて閉論理式 ξ を

(2) $\xi = \exists y\ [\ \chi(\bar{k}, y)\ \wedge\ F(y)\ ]$

と定める。このとき次の (3) (4) がなりたつ。

(3) $\vdash F(\overline{G(\xi)}) \rightarrow \xi$

(4) $\vdash \left(\neg F(\overline{G(\xi)})\right) \rightarrow \neg\xi$

実際にいま、$y = \overline{f(k)}$ とおくと

$$f(k) = G(A_k(\bar{k})) = G(\ \exists y\ [\ \chi(\bar{k}, y)\ \wedge\ F(y)\ ]\ ) = G(\xi)$$

である。(3) を示すために $F(\overline{G(\xi)})$ が、なりたっているとする。

$$F(\overline{G(\xi)}) = F(\overline{f(k)}) = F(y)$$

であって、更に (1) と $y = \overline{f(k)}$ から$\chi(\bar{k}, y)$ を得る。すなわち

$$\exists y\ [\ \chi(\bar{k}, y)\ \wedge\ F(y)\ ]$$

が成立している。(2) から ξ を得る。

したがって

$$\vdash F(\overline{G(\xi)}) \rightarrow \xi$$

である。これで (3) が得られた。

今度は (4) の前件 $\neg F(\overline{G(\xi)})$ を考える。$\overline{G(\xi)} = \overline{f(k)} = y$ であったから、この前件から $\neg F(y)$ を得る。したがって

(5) $\vdash\ \forall y\ [\ y = \overline{f(k)} \rightarrow \neg F(y)\ ]$

である。ここで既に (1) から

$$y = \overline{f(k)} \leftrightarrow \chi(\bar{k}, y)$$

が、言えているので (5) における $y = \overline{f(k)}$ を $\chi(\bar{k}, y)$ で置き換えて (6) を得る。

(6) $\vdash\ \forall y\ [\ \chi(\bar{k}, y) \rightarrow \neg F(y)\ ]$

以下 (6) を次のように同値変形する。

$$\forall y\ [\ \neg\chi(\bar{k}, y)\ \vee\ \neg F(y)\ ]$$
$$\leftrightarrow \neg\exists y\ [\ \chi(\bar{k}, y)\ \wedge\ F(y)\ ]$$
$$\leftrightarrow \neg\xi$$

したがって、¬ ξ が得られる。すなわち

$$\vdash \left(\neg F(\overline{G(\xi)})\right) \rightarrow \neg\xi$$

である。(3) (4)から、この ξ について

$$\vdash (\ \xi \leftrightarrow F(\overline{G(\xi)})\ )$$

である。【証明おわり】

3-4. 不完全性定理

定理 6 により $\vdash (\ \xi \leftrightarrow F(\overline{G(\xi)})\ )$ をみたす閉論理式 ξ の存在が言えた。そこで F として定義 2 で定めた論理式 Bew の否定を取って、自己言及的な論理式 ξ を考える。

[定義 7] ξ は、*PA* の閉論理式とする。

$$\vdash (\ \xi \leftrightarrow \neg \mathrm{Bew}(\overline{G(\xi)})\ )$$

をみたす ξ をゲーデル文という。

【定理 8】*PA* のゲーデル文 ξ について (7) (8) がなりたつ。

(7) *PA* が無矛盾ならば、*PA* のもとで ξ は証明できない。

(8) *PA* がω-無矛盾ならば、*PA* のもとで ¬ξ は証明できない。

[証明] まず (7) から証明する。そこで *PA* は無矛盾であって ⊢ξ と仮定する。つまり ξ は *PA* で証明可能である。このとき、定義 3 で述べたように $\vDash \mathrm{Bew}(\overline{G(\xi)})$ である。Σ_1-完全性により $\vdash \mathrm{Bew}(\overline{G(\xi)})$ となる。ξ はゲーデル文であるから定義 7 から ⊢¬ξ である。つまり、*PA* のもとで ξ と ¬ξ の両方が証明可能になってしまう。これは、*PA* が無矛盾であるとの仮定に反する。したがって *PA* が無矛盾ならば、ゲーデル文 ξ は *PA* のもとで証明されることはない。

つぎに (8) を示すために ⊢¬ ξ を仮定する。ξ はゲーデル文であるから 定義7から $\vdash \mathrm{Bew}(\overline{G(\xi)})$ である。*PA* については今、ω-無矛盾性が仮定されている。いま仮に ⊢ξ でない、と仮定すると *PA* のもとで ξ に至るいかなる証明も存在しない。換言すれば、どんな n に対しても $\boldsymbol{Beweis}(n, G(\xi))$ は成り立たない。このとき、表現定理によって

$$\vdash \neg \mathrm{Beweis}(\ \bar{n}\ , \overline{G(\xi)}\ )$$

が言える。 つまり、

$$(\forall n)\ \vdash \neg \mathrm{Beweis}(\ \bar{n}\ , \overline{G(\ \xi\ )}\ )$$

である。更に $\vdash \mathrm{Bew}(\overline{G(\xi)})$から

$$\vdash \exists x\ \mathrm{Beweis}(\ x\ , \overline{G(\xi)}\ )$$

を得る。しかし、これは ω-矛盾である。よって ⊢ξ でなければならない。このことは *PA* がω-無矛盾ならば ⊢ξ かつ ⊢¬ξ であることを意味する。つまり *PA* が矛盾することになる。したがって *PA* はω-矛盾する。これは不合理。したがって(8)を示すために置いた最初の仮定 ⊢¬ξ は成り立たない。つまり、*PA* のもとで ¬ξ は証明されない。【証明おわり】

定理 8 の (7)と(8) をまとめて次の定理 9（不完全性定理）が得られる。

【定理 9（Incompleteness Theorem）】自然数論を含む体系 S がω-無矛盾ならば、S の中に閉論理式 ξ が

存在して ξ は S から証明できないし、かつ ¬ξ も S から証明できない。

ゲーデル文 ξ の意味するところは、「自分自身は証明されない」という内容である。ξ は閉論理式であるから(1)真であるか、(2)偽であるかのどちらかである。
(1)　ξ が真であるとすると ξ は証明されない、という主張が正しいことになる。実際に定理 8 の (7) で示したように、ξ は証明されない。
(2)　ξ が偽であるとすると ¬ξ が真である。つまり、ξ は証明されることになる。しかし定理 8 で示された通り ξ は証明されない。したがって (2) は起こらない。つまり、(2) ではなくて (1) の方が正しいわけである。したがって、定理 8 の (7) から次の定理 10 を得る。

【定理 10】自然数論を含む体系 S が無矛盾ならば S の中に、ある解釈の下で真ではあるけれども S の中で証明することが出来ない閉論理式が存在する。

実は“体系が完全（complete）である”と言った場合に、文脈によって次の 2 つの違いがある。
(i)　意味論的完全性
不完全性定理の発表に先立つ 1 年前、すなわち 1930 年にゲーデルは 1 階述語論理の体系が完全（complete）であることを発表している。この文脈での完全の意味は、恒真である論理式は必ず証明することが出来る程度に 1 階の述語論理が完全であるという意味である。
(ii)　統語論的完全性
本稿で述べた不完全性定理における完全の意味は、任意の閉論理式 ξ について ξ かさもなくば ¬ξ のどちらか一方は証明可能であるという意味である。波線部分の否定をとると「ξ も ¬ξ も両方とも証明されないような閉論理式 ξ が存在する」という主張になって、まさしく定理 9 では、この統語論的完全性が議論になっていることがわかる。

§4.　Turing の定理に至る道筋

不完全性定理が発表されたのは 1931 年である。それ以前には、Hilbert 学派らによって無矛盾かつ完全な自然数論の公理系の構築が、計画されていた。しかし不完全性定理の主張する内容は、この計画の眼前に大きく立ちはだかってしまったわけである。それでも 1936 年には Gentzen によって、有限の立場を緩めたある立場で、自然数論の公理系の無矛盾性が証明された。つまり自然数論の中であれば、無矛盾な公理系は存在するのである。しかし、その無矛盾な自然数論の中には肯定もその否定も証明できない文が、潜んでいるわけである。Gödel のつくったいわゆるゲーデル文は、証明できるとか証明できないという内容であって、自然数に関する言明ではない。つまり超数学的な言明である。自然数に関する言明で真である（集合論的には証明できる）がペアノ算術の体系 PA では証明できない“数学の”命題がついに 1977 年になって Paris と Harrington によって発見された。1970 年には Hilbert's Tenth Problem が否定的に解決されるに至った。否定的に解決されるとは、ヒルベルトの第 10 問題が決定不能、つまり第 10 問題を一般的に解くためのアルゴリズムは、存在しないことが Matiyasevich によって証明されたのである。

4-1.　Post の対応問題

このように第二次世界大戦後、具体的な決定不能問題がいくつも発見されるに至った。決定不能問題の歴史を考える上で、Emil L. Post と Alan M. Turing の仕事は重要である。

【Post の対応問題（Post's Correspondence Problem）】*PCP*

Σ は有限個の記号からなるアルファベットである。Σ 上の記号列の全体を Σ^* であらわし、Σ^* から空列を取り除いた集合を Σ^+ であらわす。A, B は Σ^+ の記号列からなる有限集合とする。ただし A, B には同数の記号列があるものとする。すなわち、

$$A = \{w_1, w_2, \dots, w_k\}$$
$$B = \{x_1, x_2, \dots, x_k\}$$

とする。

PCP が解を持つとは、

$$w_{i1}\ w_{i2}\ \dots\ w_{im} = x_{i1}\ x_{i2}\ \dots\ x_{im}$$

を満たす整数 $m \geq 1$ と整数の列 $i_1, i_2, \dots, i_m$ が存在することである。長さ m の整数の列 $i_1, i_2, \dots, i_m$ を 1 つの解という。

Post の対応問題とは、与えられた 2 つの有限集合 A, B について、*PCP* が解を持つか否かを決定するアルゴリズムを求める問題である。Turing 機械の停止問題は、Post の対応問題に還元可能である。したがって Post の対応問題が可解であるとすると、Turing 機械の停止問題が可解になってしまう。これは不合理である。そのような訳で、Post の対応問題が非可解であることが示される。

このように Turing 機械の停止問題が帰納的に可解

でないという主張は、多くの決定問題の否定的解決のために極めて重要である。

4-2.　Turing 機械

§3 で述べた不完全性定理は“証明”に関する議論であった。証明自体は定義 1 で定められている通り、論理式の有限列である。n が非常に大きければ、長い証明ということになる。人工知能にせよ数理知能にせよ“知能”を特徴付けるものは、推論である。定義 1 では論理式　A から論理式 B が推論されることを $A \vdash B$ で表した。beweisbar（provable：証明可能）を表すこの記号　$\vdash$ を computable（計算可能）に置き換えて考えてみよう[5]。すなわち、表 6 のような置き換えを考える。

表 6　証明から計算への置き換え

論理式 $A, B, C, \ldots$	⇒	語 $\alpha, \beta, \gamma, \ldots$
$A \vdash B$ (B は A から証明可能)	⇒	$\alpha \vdash \beta$ (β は α から計算可能)
証明図	⇒	計算

A.M. Turing と E.L. Post は、それぞれ独立に、しかも同時期に証明論で得られた定式化を計算論として定式化した。目的は不完全性定理の証明の中に登場した原始帰納的関数を一般化した Herbrand-Gödel-Kleene による帰納的関数の概念のそれぞれのアイディアによる具体化である。Turing によって得られた計算システムは、Turing 機械（Turing Machine）と呼ばれており、Post による計算システムは双正規系（Binormality）と呼ばれている。本稿では、アルゴリズムの定式化の仕組みとして Turing 機械を採用する。

4-3.　万能 Turing 機械

任意に部分帰納的関数 $f(x)$ が与えられたとき、Kleene の標準形定理によって

(1)　$f(x) \simeq U\ (\mu y\ \ T\ (e, x, y)\ )$

を満たす自然数 e が存在する。この自然数 e は f に依存して決まることから f のゲーデル数と呼ばれる。(1) の右辺に注目する。$U, \mu y,\ T$ は固定的である。それに対して e と x だけは、左辺に指定された f と x に従って変化する。このことから、つぎのような仕組みを考えることができる。

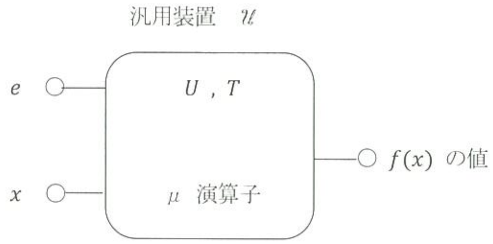

図 7　　プログラム内蔵方式のヒントとなった万能 Turing 機械

図 7 の汎用装置 $\mathcal{U}$ を万能 Turing 機械と呼ぶ。今日のことばで言えば e はプログラム、x は入力データ、$\mathcal{U}$ はそのアーキテクチャをプログラム内蔵方式とするコンピュータである。実際に万能 Turing 機械 $\mathcal{U}$ は、テープ上に入力 x とプログラム e を配置し、e から処理の仕方を教わりながら計算を進め最終的に $f(x)$ を出力するのである。一方、プログラム内蔵方式とは別に ROM 方式というアーキテクチャもある。自動販売機やインヴェーダ・ゲームのようなハードウェアとソフトウェアが一体化している機械（ファームウェア）がそれである。1 台の Turing 機械 M はファームウェア化されており、5 項系列の有限集合で定義される。5 項系列は、記号の有限列であるから 1 台の Turing 機械は、§3 で述べたゲーデル数の定義の (iii) における算術化によって 1 つの自然数 e に凝縮させることができる。この自然数 e を ROM（Read Only Memory）化させて、1 台の Turing 機械 M が実現されるわけである。M に入力 x を与えたとき、M が正常に終了してテープ上に計算結果 $M(x)$ が得られたとする。M が帰納的関数 f を計算する Turing 機械で、M に入力 x が与えられたとき

$$M(x) = U(e, x) = f(x)$$

が成り立つ。

万能 Turing 機械は f のゲーデル数 e に従い、M の動作を模倣するわけである。したがって次の関数 $rechnen(p, x)$ を計算する万能 Turing 機械が存在する。

---

[5] “provable” を ”computable” と同一視する発想は、すでに Gödel による不完全定理の証明の中にある。不完全定理の証明の中で議論されている形式的体系に求められる条件は無矛盾であること、算術を含むこと、それから帰納的であること、すなわち勝手に与えられた閉論理式がその体系の公理であるか否かが有限の手続きで決定できるという性質の 3 つが求められる。欲をいえば統語論的完全性も欲しいとろであるが、上記の 3 つと統語論的完全性は両立し得ないというのが、不完全性定理のいうところである。

$$rechnen(p,x) = \begin{cases} M(x) & p \text{ が Turing 機械 } M \text{ のゲーデル数になっているとき} \\ 0 & p \text{ が Turing 機械のゲーデル数には、なっていないとき} \end{cases}$$

入力 x に対して常に $M(x)$ が値を持つとは限らない。そのような場合、$rechnen(p,x)$ を計算する万能 Turing 機械は停止しない。この不気味さを解消するために全域関数 $ganze\text{-}rechnen(p,x)$ を定める。

$$ganze\text{-}rechnen(p,x) = \begin{cases} M(x) & p \text{ が Turing 機械 } M \text{ のゲーデル数になっていて、しかも } M(x) \text{ の計算が正常に終了するとき} \\ 0 & \text{そうでないとき} \end{cases}$$

【定理 11】 $ganze\text{-}rechnen$ は、計算可能ではない[6]。

[証明] 背理法で証明する。$ganze\text{-}rechnen$ が計算可能、すなわち $ganze\text{-}rechnen$ を計算する Turing 機械 V があったと仮定する。つぎの関数 $quit$ を考える。

$$quit(x) = \begin{cases} 1 & ganze\text{-}rechnen(x,x) = 0 \text{ のとき} \\ 0 & ganze\text{-}rechnen(x,x) > 0 \text{ のとき} \end{cases}$$

$quit$ の定義から、すべての x に対して

(*) $quit(x) \neq ganze\text{-}rechnen(x,x)$

である。 $ganze\text{-}rechnen$ を計算する Turing 機械 V を加工して $quit$ を計算する Turing 機械 H をつくることが出来る。この Turing 機械 H のゲーデル数を p_h とおく。 p_h を万能 Turing 機械 $\mathcal{U}$ に入れて $rechnen(p_h,x)$ を計算することを考える。 $quit$ は、すべての x に対して定義されていることから

$quit(x) = rechnen(p_h,x) = ganze\text{-}rechnen(p_h,x)$

である。特に、 x として $x = p_h$ をとったときは

(**) $quit(p_h) = ganze\text{-}rechnen(p_h,p_h)$

となって(*)と矛盾する。

したがって $ganze\text{-}rechnen$ は、計算可能ではない。【証明おわり】

つぎの述語 $halt$ を考える。

$$halt(p,x) = \begin{cases} 1 & p \text{ が Turing 機械 } M \text{ のゲーデル数になっていて、しかも } M(x) \text{ の計算が正常に終了するとき} \\ 0 & \text{そうでないとき} \end{cases}$$

【定理 12】 $halt$ は、計算可能ではない。

[証明] もしも $halt$ が計算可能であるとすると、$halt$ を計算する Turing 機械 M から $ganze\text{-}rechnen$ を計算する Turing 機械を次のようにしてつくることが出来る。

$$ganze\text{-}rechnen(p,x) = \begin{cases} M(x) & halt(p,x) = 1 \text{ のとき} \\ 0 & halt(p,x) = 0 \text{ のとき} \end{cases}$$

$halt$ が計算可能と仮定されているので、$ganze\text{-}rechnen$ も計算可能である。しかし、既に定理 11 で $ganze\text{-}rechnen$ は、計算不能であることが示されている。よって、$halt$ は、計算可能ではない。【証明おわり】

【Turing 機械の停止問題】 1 台の Turing 機械 M 及び入力 x が任意に与えられたとする。M のゲーデル数 p を算出して入力 x を $halt(p,x)$ に与えたとき、定理 12 の主張は $M(x)$ の計算が正常に終了するか否かは、帰納的に決定不能であるということである。

6 定理 11 と定理 12 の定式化は、渡辺・米崎[8]に依拠している。

$M(x)$ の計算が正常に終了しない場合、すなわち無限に計算を続けるかまたは $M(x)$ の値が確定する前に停止してしまう場合は、M が異常終了する（abort）という。

ある問題 S が与えられたとき、任意の x に対して $x \in S$ であるかそれとも $x \notin S$ であるかを決定するTuring機械 F が存在して、任意の入力 x に対して必ず F が停止するならば、S は帰納的（recursive）であると言われる。しかし問題 T の中には、適当な帰納的関数 f について

$$T = f(\omega) = \{f(0), f(1), f(2), \dots\}$$

となるものもある。このような T は帰納的可算（recursively enumerable）な問題と言われる。Turing機械の停止問題 $halt(p,x)$ は、帰納的でない帰納的可算問題である。我々は §0 の図2において人工知能で解ける問題群と数理知能で解ける問題群（＝帰納的に可解な問題群）との関係を確認している。Turing機械の停止問題 $halt$ は、図8に示すように帰納的に可解な問題のクラスの補集合に属するものとして位置づけられるわけである。与えられたアルゴリズムらしき手続き h に入力 x を与えたとき、$h(x)$ が計算結果を出して停止するのか、それとも永遠に計算を続けるのかを判定するアルゴリズムが存在しないことが証明されたわけであるから、§1 で述べた以下の本稿の主題と主張が確認されるわけである。

【本稿の主題】与えられた手続きが正しいアルゴリズムであるか否かを判定する人工知能は存在しない。

どんなロボットも人工知能を搭載しているので、この主題からつぎの主張を得る。

【本稿の主張】正しく作動する人工知能を自律的に作るロボットは、存在し得ない。

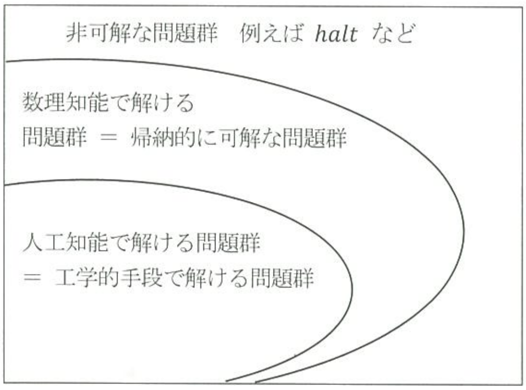

図 8　非可解問題の位置付け

§5.　おわりに

人工知能の飛躍的発展の影響や問題点を危惧する考えには、２つの種類があるようである。ひとつは、工学の進歩によってロボットや人工知能が進歩し、ホワイトカラーといわれる人々が従事して来た仕事が、高度な人工知能にとって替わられる事態を危惧する考えである。もうひとつは、人工知能が指数関数的に進歩し人間の知能をいつかは超えてしまい、最終的に人工知能が全地球人の政治や経済を支配・コントロールする時代がやって来るという考えである。このような深刻な事態を技術的シンギュラリティ（Technological Singularity）と呼ぶ。このシンギュラリティはそう遠くない日、一般的に言われているのは 2045 年であるが、意外と早くやって来るという考えである[7]。

前者については、これは正しい予測であると筆者は考えている。確かに、工学の進歩は時間軸に対して決して線形ではない。特に、人工知能の技術は進歩と停滞を繰り返して来た。21 世紀初頭の現在、人工知能技術を急速に発展させる工学的基盤や社会的基盤は整っている。したがって、これからの時代を牽引する学生諸君はどのような産業がこれから益々、発展し、どのような産業が衰退して行くのかという問題に絶えず、注意を向けるべきである [8]。一方、後者についての予測は前述したとおり、間違いである。§4 で述べて来たとおり人工知能がどのように進歩したとしても、正しく作動して更に無限ループに落ち入るかそれとも停止するかどうかを判断してくれる人工知能、所謂、全域正当性を判断してくれる人工知能は存在し得ないわけである。したがって人工知能で出来た絶対者が全地球人の政治・経済や教育を支配・コントロールする時代は、やって来ない。つまり人知を超える万能の神を、わたくし達人間は工学的につくり出すことは出来ないのである。正しさの検証は、わたくし達人間自身の手で行わなくてはならない。そのために人間は、道具を開発して来た。コンピュータや人工知能もその種の道具である。人工知能を動かすプログラムの部分正当性を検証する技術は、すでに多く開発されている [9]。つまり、プログラムにある種の制限を加えれば、そのプログラムが停止するか否かは判定可能になる。プログラムの正当性を検証する技術に関する論考については、

---

[7] 理論物理学と天文学の分野で偉大な貢献をしたスティーブン・ホーキング博士の危惧が、以下の Web ページにある。
http://www.huffingtonpost.jp/2014/12/03/stephen-hawking-ai-spell-the-end-_n_6266236.html

本稿とは別に稿を改めて考察することにしたい。

いわゆる 2045 年問題のような、人類の知性を超える大いなる超越者の到来を畏怖する心情は旧約聖書の成立した時代から存在して来た[10]。事実、ダヴィデがサターンに唆されてイスラエルの民の数をカウントし終えた後、ダヴィデは神の怒りに遭遇してしまうのである。今日、人口統計を正確に把握し、経済政策の策定に資する行為は当然のように実施されている[11]。自然現象や社会現象を数値を使って記述したり、数理モデルを構築して近未来を予測する技術は近現代に至ってはじめて行われている所作に過ぎない。長い人類の歴史の中では、数を以て合理的に現象を記述しようとするデカルト的精神の発露は、まだ未熟なのである。実際に 18 世紀の英国議会では、上記の旧約聖書の記述から人口調査の実施に反対する議員が多数、いたそうである[12]。つまり、数を以て人口を数えるような行為は人間は行ってはならなかったのである。このような所作は、神のみに許されていたのであった。

前述の通り、万能のロボットが出現して次から次へと優秀なロボットを生産するなどという時代は幸か不幸か、決してやって来ない。この事実を知って、ふと安堵に胸をなでおろす人も居るかも知れない。人類の知能を超える人工知能の出現を私たちが畏怖するのは、いわば人間の持つ本能的後ろめたさに起因するものであろう。筆者は、このような後ろめたさは失ってはいけない、デカルト的合理精神は決して万能ではない、と考えている。なぜなら、コンピュータを誕生させた人間の行う愚直とも言える合理的精神行為は、137 億年前にこの物理世界を誕生させた創造主の合理的精神行為には遠く及ばないからである。

なお、本稿で言及した Gödel と Turing の功績により理論計算機科学における顕著な業績に対しては、それぞれゲーデル賞、チューリング賞という褒彰が設けられていることを申し添える。

謝辞　法政大学名誉教授　田中尚夫先生には、執筆の段階で本稿を注意深く読んでくださいました。更に、貴重なご意見を頂戴することができました。田中尚夫先生に深謝の意を表します。

註

1) チューリング・テストについては、岩波書店刊『情報科学辞典』 p. 467 によると以下のように述べられている。「1つの部屋に計算機, 別の部屋に人をおき, どちらの部屋に計算機が入っているか知らない人が, 外から通信線を通じてそれぞれの部屋に種々の質問をしたとする. それらの質問に対する応答がいずれも適切であって, 計算機のものか人間のものか判定しにくいときに, 計算機は人間のように考えているとチューリングは定義した.」

2) “数理知能”は、不完全性定理を議論するときによく使われる“理性”と同義語である。たとえば、吉永良正著『ゲーデル・不完全性定理　……理性の限界の発見……』講談社（ブルーバックス）における“理性”と同義語である。

3) 倉田令二朗著『入門数学基礎論』河合文化教育研究所, 136 ページでは“帰納的関数＝アルゴリズムの存在する関数”という Church の提唱をいわば“「ホモサイエンスの計算能力」に対する仮説といった性格のものである”と述べている。

4) 逆理（paradox）の発見は、“数学の危機”とも言わせしめた程に深刻な問題であった。どのようにしたら逆理から逃れることが出来るのか、について多くの数学者の研究がある。最近、Hermann Weyl, “Das Kontinuum” の訳本が、田中尚夫・渕野　昌によって『ヘルマン・ヴァイル　連続体　解析学の基礎についての批判的研究』日本評論社（2016 年 2 月 20 日）として出版された。

5) Principia Mathematica の日本語訳については、岡本賢吾・戸田山和久・加地大介訳『プリンキピア　マテマティカ　序論』哲学書房（1988 年）が刊行されている。

6) 金山　裕著『プログラム理論』法政大学工学部 講義ノートより引用

7) 田中一之編『ゲーデルと２０世紀の論理学　第３巻　不完全性定理と算術の体系』の p. 78 において著者（鹿島　亮）はゲーデル数について、以下のように記している。「現実のコンピュータの動作原理では, プログラムを数値データ（２進数）で表現してそれを計算対象として扱う（たとえばコンパイラというプログラムは, プログラムを計算対象として作動する）ことが重要であるが, この考えはまさに「証明」をゲーデル数で表現してそれを体系内で証明対象として扱うことと同じである. その意味でも（第１章でも述べたが）ゲーデルの不完全性定理はその後のコンピュータの理論の先取りをしているのである.」

8) AERA（アエラ）No. 26 号, 朝日新聞出版（2015 年 6 月 15 日号）pp. 12-13 には、今後 10～20 年後になくなる仕事と残る仕事のリストが掲載されている。この記事の原典は、オックスフォード大

学オズボーン准教授による論文『雇用の未来』と記されている。

9) たとえば、C. A. R. Hoare. “An axiomatic basis for computer programming”, Communications of the ACM, 12(10), p.p. 576-580, 1969 や http://atva2016.gforge.inria.fr/ など。

10) 筆者が、この旧約聖書の記述を知ったのは、新井紀子・新井敏康著『計算とは何か』p.153 による。

11) 米国で行われた世界ではじめての国勢調査に、IBM 社の Punched Card System が貢献した。Punched Card System を開発したホレリス博士が旧約聖書の記述をどのように考えていたかについては手元に資料がない。

12) この史実も、新井紀子・新井敏康著『計算とは何か』p.153 の脚注に記述がある。

参考文献

[1] A.M. Turing, “*ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHEIDUNGSPROBLEM*”, Proc. London Math. Soc., Ser.2, Vol. 42(1936-37), pp. 230-265.

[2] K. Gödel, “*Über formal unentscheidbare Sätze der Principia matehmatica und verwandter Systeme I*”, Monatshefte Math. Phys., 38(1931), pp. 173-198.

[3] 田中一之・鹿島 亮・角田法也・菊池 誠著『数学基礎論講義 不完全性定理とその発展』日本評論社, 1997 年

[4] 田中一之編『ゲーデルと２０世紀の論理学 第２巻 完全性定理とモデル理論』東京大学出版会, 2007 年

[5] 田中一之編『ゲーデルと２０世紀の論理学 第３巻 不完全性定理と算術の体系』東京大学出版会, 2007 年

[6] 林 晋・八杉満利子訳・解説『ゲーデル 不完全性定理』岩波文庫, 2015 年

[7] 前原昭二著『数学基礎論』朝倉書店, 昭和 53 年, pp. 117-118.

[8] 渡辺 治・米崎直樹著『計算論』日本評論社, 1997 年, pp. 66-72.