

# 「コンタクト・パラドックス」とその同類たち

稲葉 振一郎

## 1. イントロダクション

現代のSETI (Search for Extraterrestrial Intelligence) においては、METI (Messaging to Extraterrestrial Intelligence, 能動的 (Active) SETIともいう) の是非を巡って論争がある。初期のオズマ計画以来、単なるSETI, ETIからのメッセージの受信やその存在 (の兆候, 痕跡) の探索にとどまらず、存在しているかもしれないETI, ETC (Extraterrestrial Civilization) に対して、地球人類の側から積極的にメッセージを発信することの是非——公共的インパクトを持ちうるそうした事業を研究者が勝手に行うことの是非、あるいは発信それ自体の是非——はしばしば問われてきたが、とくに近年ではMETI積極派と消極派・否定派との間での対立、論争が先鋭になりつつある<sup>(1)</sup>。Cooper (2019 = 2021) によれば、その中で「コンタクト・パラドックス」とでも呼ぶべき問題が浮上してきている。

もちろんSETIそれ自体が、数十年の歴史を重ねながらも実証的には未だ成果を上げておらず、理論的な観点からもETIとのコンタクトの可能性それ自体への疑義が有力となっている<sup>(2)</sup>以上、METIの是非を問う以前にそもそもSETIに対しても、とりわけ実証的な試みについてはその是非が問われても仕方がないのが現状ではある。ただ問題の性質上、コンタクト可能性への懐疑論も「ETIは存在しない」「仮にETIが存在していたとしても、それとのコンタクトは不可能である」、という主張ではありえず、「ETIが存在する可能性は極めて低い」「仮にETIが存在していたとしても、それとのコンタクトの可能性は極めて低

い」というものでしかありえない。それゆえにあくまでもSETI自体の意義を肯定した上での、METIそれ自体の意義、是非を問う議論の意味もまた、消滅するものではない。のみならず、以下に見ていくように、人工知能開発や宇宙植民など、遠未来の人類の末裔が直面するかもしれない多くの課題において、この「コンタクト・パラドックス」と似通った構造を持つ、ある種「パラドックス」と呼びうるような問題群が浮上してくると思われる。それらが共有する構造自体は一定の普遍性を持つと考えられるため、その観点からも掘り下げるに値する。本稿ではまずMETIを巡る消極論と積極論の対立を概観し、「コンタクト・パラドックス」の構造を明らかにした上で、宇宙倫理学やその他の領域における類似の諸問題と比較する。その上でそれらに共通するある構造について指摘し、その脱パラドックス化の可能性を示唆する。

## 2. METIを巡る論争

METI消極論、すなわち、受動的な探索にとどまらない能動的なSETI、METIが当面の人類、つまり観測技術はある程度発展させて、太陽系外のみならず宇宙全体についての情報を収集してはいるが、居住拠点はあくまでも出身惑星たる地球にとどまっており、工学的な環境改造も宇宙においてはまだ行っていないような技術水準にある現在並びに近い未来における人類にとってリスクであり、回避されるべき、少なくとも慎重であるべきである、という議論はいまのところ優勢であると言えよう。その理由について、この論争を瞥見した一般書であるCooper (2019=2021)、宇宙倫理学教科書であるGreen (2021)を踏まえつつ考えてみよう。

仮にコンタクト（単なる発見にとどまらず、双方向的なコミュニケーションの開始）が現実的に可能な範囲に（具体的には例えば、電波などによる双方向通信が、人類にとっての一世代かせいぜい二、三代程度ペースで可能な、

数光年からせいぜい数十光年の距離に) ETI, ETCが存在しているとする。ところが、我々がそのように比較的容易にコンタクト可能な、それゆえそれとのコンタクトの可能性について真剣に考えなければいけないようなETIの文明・技術水準は、我々のそれを優に凌いでいる可能性が高い。なぜか？

第一に、そもそも確率の問題として、それほどの近距離に、しかも我々とそれほど技術水準の変わらない(具体的には、宇宙観測の能力はあるが、身を以ての大々的な宇宙進出はまだこれからという程度の) ETI, ETCが存在しているという可能性自体が非常に低い、ということに注意せねばならない。我々の技術文明が激烈な変化を経験した数十年、数百年はもちろんのこと、農耕開始以来を目安とした1万年というタイムスケールでさえ、宇宙的観点からはほんの一瞬である。仮に今現在、我々と同時代にETI, ETCが存在していたとしても、その技術水準は現在の我々のそれと大きくかけ離れている、すなわち我々を遙かにしのいで、既に星間文明を実現しているか、あるいは逆に、母星を飛び出すどころか、なお産業化以前にとどまっているか、のどちらかである可能性が高い。

その上で第二に、現代の我々の探査能力、更にはそれを支える技術水準や社会的機構に鑑みたとき、我々のそれと同水準か、あるいはそれ以下の文明・技術水準のETIとコンタクトできる可能性はそもそも低い。太陽系から数光年からせいぜい数十光年という比較的狭い範囲内においても、そして20世紀末以降の系外惑星の発見ラッシュ、更に今後数十年のうちに期待される探査能力の向上を計算に入れてもなお、地球・太陽系からの光学・電波観測に頼っての受動的SETIによっては、先方が意図的に発信した信号の受信の場合を除けば、特定の惑星上にその活動範囲が限定されているETI, ETCをもっぱらこちらからの観測によって発見することは非常に困難であると思われる。地球・太陽系からの観測によってなしうることは、天体のスペクトル観測などを通じてバイオシグネチャーやテクノシグネチャーを発見し、そこに生命や技術文明の存在す

る可能性を確率的に見積もることができる程度であろう。だとすればそのような探査によって発見可能なETI, ETCは、遠方からでも容易に観測できるようなテクノシグネチャーを伴う大規模な宇宙工学を実現した高度な技術文明である可能性が高い。あるいは向こうからこちらを発見し、信号や場合によっては探査機、宇宙船を送り込んでくるような高度なETI, ETCならばなおさらのことである。

さて以上の推論<sup>(3)</sup>が正しく、我々が探知し、更にコンタクトを取ることができるようなETI, ETCはほぼ確実に我々より圧倒的に優位な技術水準にあるとしたら、そのような相手とのコンタクト・交流は我々人類にとって非常にリスクである、となぜ言えるのだろうか？

第一に、仮に相手が悪意・害意をもって人類に接触し、人類文明を攻撃、侵略、あるいは支配、搾取、最悪の場合殲滅しようとしたときに、人類がこれによく対抗して自衛できる可能性が低い、と思われる。なぜなら相手の技術水準は人類のそれより圧倒的に高いはずだからである。

第二に、仮に相手に悪意がない、あるいはむしろ善意でコンタクト・交流してきた場合にも、未知の、あるいは想定外のアクシデントが起きる可能性は決して無視できない。

これらの主張は人類史の経験からの類推によっても支持できるだろう。無数の他の生物を絶滅に追いやったことを差し引いて、人類の範囲内に話を限定したとしても、生産力・軍事力によって優越した文明が他の文明を侵略し、征服して衰亡させた例は人類史において枚挙に暇がない。20世紀後半、そうした歴史に対する反省を経て以降、文化接触がより注意深く、デリケートに行われるようになってからも、少数派の文化コミュニティが、より大規模で生産力の高いグローバル社会との接触によってスポイルされるという問題は深刻である<sup>(4)</sup>。

さて、以上のような消極論、悲観論に対して、能動的SETI、METIについての楽観論・肯定論者の見解はどのようなものだろうか？ 大雑把に言えばそれは「高度文明の発展にとっての利他主義の合理性・宇宙的普遍性」仮説とでもいうべきものへの信頼に立脚しているように思われる<sup>(5)</sup>。それによると、我々人類とコンタクトしうるような（我々のメッセージを受信しそれに返答しうるような、あるいは向こうからこちらまで来るような）ETCは非常に高度であり、現在我々が直面しているようなグローバル・リスク、ニック・ポストロムの言うExistential Risk——環境問題や政治対立などの内的リスクのみならず、地球科学的・天文学的破局のような外的危険も含む——<sup>(6)</sup>の多くを既に克服している可能性が高い。とりわけ、グローバル文明を脅かすような過度な攻撃性は克服しているであろう。そのような文明であれば他の知性、他の文明に対しても侵略的である可能性は低い。

それに対してこうした楽観論に対する批判としては、以下のようなものが考えられる。

第一に、仮に高度な宇宙文明を建設したETI、ETCが仲間内での過度な攻撃性を克服していたとしても、それが他者、他種の知的生命、文明に対しても同様である保障はない。自己の安全を確保するために、他の文明に対して攻撃的に振る舞う可能性もある。

第二に、仮に先の宇宙的利他主義仮説が正しかったとしても、既に見たとおり、善意や害意の不在が無害を意味するわけでは必ずしもない。たとえ平和的にはあれ、優越した宇宙文明との交流によって地球人類の文明が何らかの悪影響を受ける可能性は無視できない。

### 3. コンタクト・パラドックス

以上のように、Cooper (2019=2021)、Green (2021) の整理を踏まえるな

らば、現状ではMETI、更に言えばコンタクトそのものに対しての否定論、とは言わないまでも消極論の方が有力であるようだ。このような論争状況を踏まえて、キース・クーパーは「コンタクト・パラドックス」とでもいうべき問題があると主張する (Cooper (2019=2021))。すなわちそれは「人類がETIとのコンタクトを安全に行い有益な成果を得るための必要条件を、他ならぬETIとのコンタクトの経験なしには得ることが難しい」というものである。

これが「コンタクトのパラドックス」だ、われわれは、ほかの生命とコンタクトをとる地球外生命の証拠を宇宙に探し求めている。人類がこの宇宙で他者と共存できるように。ところがわれわれは、コンタクトをとろうとすべきかどうかの判断に自信がない状況にある。利他行動の仮定から、生命がもちうる知能のタイプの違いまで、天の川銀河に入植する能力から、フェルミのパラドックスの謎まで。これまでSETIについて明らかにした何もかもが、そして言うまでもなく、われわれ自身の歴史における偶然のコンタクトさえもが、用心せよとわれわれに教えている。「宇宙に私たちより進んだ種族がいて、彼らがじっと身をひそめていたら、私たちが知らないことを彼らが知っていると考えてはいけなんでしょうか？」とデイヴィッド・ブリンは問うている。コンタクトは危険なことであり、静かでひょっとしたら危険かもしれないジャングルに、何が棲んでいるかを知らずに大声を響かせるのは、賢明ではない可能性があるのだ。(Cooper (2019) p.183. = (2021) 334頁。)

「パラドックス」という語で何を意味するかは文脈によって多様であり、この場合はもちろん明確な論理的矛盾や不整合を意味しているわけではないので、「必要条件」ももちろん厳密な意味での論理的なそれではないことに注意すれば、クーパーの指摘それ自体は興味深く、ある程度の普遍性を有すると思われるので、ここで問題の構造を我々なりに定式化しよう。

実証研究への公的予算投入の規模などをめぐる極端に実際的なレベルを除けば、広義の（主として受動的な）SETIそれ自体の是非についてはさほどの論争はない。すなわち、一部の研究者が独自の問題関心に基づき、自ら調達できる資源の許容する範囲においてSETIを研究することの意義を否定する意見はほとんどない。そもそも「フェルミ・パラドックス」の題目のもとで提示される現代のSETI懐疑論自体が、主として天文学、宇宙論、宇宙生物学等の理論と（非SETI的な）観測の成果に立脚したものではあれ、広義のSETI研究の成果である、とさえ言える。では何が疑問に付されているのか？ 観測、受信を通じての受動的SETIそれ自体については、懐疑論において提示されているような「費用対効果から見てあまり有意義ではない」という批判はあるが、それ自体として「すべきではない」という強い禁止論は管見の限りでは少ない<sup>(7)</sup>。それに対して興味深いのは、積極的、能動的にETIに物理的に接触する、あるいは遠隔においてでも双方向的なコミュニケーションをとることを目指すMETIについては、ただ単に「費用対効果的にみのりがあるとは思われない」というのではなく、むしろその成功の可能性を肯定的に評価した上での、成功それ自体がネガティブな効果をもたらす可能性に注目して、積極的に「すべきではない」との批判、禁止論が持ち上がっているということである。逆に言えば、狭義の、受動的なSETIについては「成功の見込みはあるが、その成功がネガティブな帰結をもたらす恐れがあるのでやめたほうがよい」という類の議論は珍しい<sup>(8)</sup>。

ここまでの議論を整理してみると、以下のような立場を区別できる。

- a. （受動的）SETIが成果を上げる可能性は無視できるから、そこに労力を割いても虚しい。
- b. （受動的）SETIが成果を上げる可能性は無視できないから、そこに労力を割く甲斐がある。

- c. (受動的) SETIが成果を上げる可能性は無視できないからこそ、そのリスクに鑑みて慎重になるべきである。
- d. METIが成果を上げる可能性は無視できるから、そこに労力を割いても虚しい。
- e. METI が成果を上げる可能性は無視できないから、そこに労力を割く甲斐がある。
- f. METIが成果を上げる可能性は無視できないからこそ、そのリスクに鑑みて慎重になるべきである。

ここで通常の(受動的)SETI懐疑論はaにあたり、通常の(受動的)SETI積極論はbに当たるが、cの例はあまり見かけない。ただ単にETIの存在が知られること自体が、人類にとってリスクとなりうるとはほとんど思われていないということであろう<sup>(9)</sup>。dが独立に論じられる例はほぼ見かけないが、aは当然にこれを含意する(aからdが帰結する)と考えられる。そしてeがMETI積極論、fがMETI懐疑論、というよりMETI慎重論・否定論ということになる。

更にこれらの立場は互いに排他的というわけでは必ずしもない。e、fはbを前提しているであろう。もちろんfの立場を取るがゆえに、観測についてもそれが先方に逆探知の可能性を開くリスクを鑑みた慎重論を取ることがありうる<sup>(10)</sup>が、fは基本的には、コンタクトのリスクを減らすためにこそ、もしそれが可能であれば先方に知られることなく一方的にこちらがETI、ETCを探知し観測することを望む立場であろう。またdは上述の通りaから帰結する主張だが、論理的に言えばbと矛盾するわけではない。というよりaをとるならばわざわざdを独立の主張として立てる理由はほとんどない。

以上については、意思決定理論の枠組みを使って帰結主義的にETI、ETCとのコンタクトの利害得失についての評価を試みると少し見通しがよくなる。詳



細は補論に譲るが、要約すると、

- ・ETIとのコンタクト（単なるSETIにとどまらず、METIをも含めた双方向的コミュニケーションの成功）から得られるであろう利益を最大化する、あるいはリスクを最小化することを目指すという観点からすれば、コンタクトに費やすべき努力の最適水準とでもいうべきものが存在し、それは想定次第ではゼロになりうる。
- ・更にその最適水準を考えるとときに重要であるのは、 $\alpha$  「どれだけの資源・労力を投入すればコンタクトの成功率がどれくらい上がるのか」だけではなく、 $\beta$  「コンタクトが成功したとして、その結果どのような利益がもたらされるのか、あるいは危険が生じるのか」である。
- ・おそらく $\alpha$ と $\beta$ は独立である、つまり人類がコンタクトにかけた努力と、そこから得られる利益、あるいは損害とは無関係であると考えたほうがよい。
- ・ $\alpha$ が具体的にどのような形状の関数になるのか、は不明だが、おそらくは大きな努力を投じればそれだけ成功率が高まる、という増加関数であろう。しかし $\beta$ がどのような確率分布として考えられるのか、については、我々は多くを知らない。それはそもそも人類にとって未経験の事象なので、経験から学ぶ、過去の観測例の頻度分布から外挿するという手は使えず、別の事例からのアナロジー、またそうしたアナロジーを正当化する理論的考察で対応するしかない。

このように整理してみると、クーパーのいう「コンタクト・パラドックス」は、論理的な意味での、たとえば「一見妥当な前提からの論理的帰結が、前提と矛盾し、それと両立しない（ように見える）」とか、あるいは「それぞれは妥当に見える一連の命題が両立しない（ように見える）」といったパラドックスではない（それは「フェルミ・パラドックス」についても同様である）。「コンタ

クトを成功裡に、つまりそのリスクを回避し、あるいは抑制し、そこからプラスの利益を引き出すような形で行うための必要条件のひとつに、コンタクトから得られる損益についての確率分布の知識があるとしたら、そのような知識は、決して事前には、つまり実際にコンタクトを、しかも複数回経験することなしには得られない」というのがその含意である。我々は上の $\beta$ について事前に推測するしかなく、そのために使える知識があまりにも少ないのだ。もちろんそれは決してコンタクトの不可能なることを意味してはいない。ただ、コンタクトを安全に成功させることができる見込みが非常に低い、ということである。実行不可能という意味ではなく、実行することに大きな危険が伴う、ということだ。

もちろんここでの「そのような（より具体的に言えば $\beta$ についての）知識は、決して事前には、つまり実際にコンタクトを、しかも複数回経験することなしには得られない」という想定に疑問を差し挟むことはできる。我々はETIとのコンタクトの可能性について考えるに際して全く無知であるわけではない。ただし我々はそこで根拠のない（より正確には、その根拠を我々が理解していない）経験則に頼ることはできず、科学的な根拠を持った理論的予測を主たる頼みとするしかないのである。たとえば先に我々は人類史における文化接触・文明間交流の経験から得られた知見が、ETI、ETCとのコンタクト、交流において何が起りうるかを考えるために役に立つ、との前提で若干論じたが、もちろん厳密には無反省にこのようなアナロジーを適用すべきではなく、人類内、同種内でのコミュニティ間衝突と、恒星間文明衝突との間にアナロジーが成り立つことの正当化を行わなければならない。つまり人類内での紛争、また行動生態学などにおける種間・種内競争についてのデータの蓄積から見出された分布を単純に外挿するだけではだめで、それらのデータを適切に説明できる理論モデルを選び出して、それを適切に再解釈して恒星間文明衝突の理論モデルを作り出さねばならない。これは困難ではあるが、全く不可能というわけではな

いだろう<sup>(11)</sup>。それゆえ、クーパーのいう「コンタクト・パラドックス」が「コンタクトの成功のために必要な条件は、コンタクトの成功それ自体を通じてしか得られない」ということを意味するのであれば、それは厳密に言えば誤りであり、実際には状況はそこまで絶望的というわけではない。

とは言え当然我々は恒星間文明衝突の具体的事例を持ち合わせていないため、モデル構築にあたってどのような要因を組み込まねばならないかについて不十分な知識しか持たない。ETI, ETCであっても同じ宇宙の中に存在し、同じ物理法則に従っているのであれば、我々が知っている政治学、経済学、心理学、生物学の形式的理論によってその振る舞いはある程度理解し予測することさえできよう。しかしもちろん我々はETI, ETCが具体的にはどのような素材でできあがっており、どのような歴史的経過を辿って進化してきたのかは当然知らない。それらはおそらくダーウィンのな進化を経てきた生物（ないしそれに由来する機械）であろうが、その進化プロセスを具体的に担った物理メカニズムは、（たとえタンパク質ベースであったとしても）地球上の核酸ベースのものとはおそらく異質だろう。またその行動主体たちの振る舞いはゲーム理論によって解釈可能という意味では合理的であろうが、そのような合理性を具体化する実際のメカニズムが、我々の知る脳神経システムと全く同質ということはあるそうにない。現代の我々人類を対象とした行動経済学の知見は、形式科学としてのゲーム理論が予想する理想的な合理性と、実際の人間の行動メカニズムが、おそらくは人間の生物学的・物理化学的性質のおかげでシステムティックにずれていることを発見している<sup>(12)</sup>。そしておそらくはETIたちの場合にも、それぞれの進化生物学的、歴史的個性によって、理想的な合理性からのズレが見られるだろう。しかしそのズレ方が我々人類のそれと全く同じである保証もまたない。すなわち、我々はETI, ETCとの紛争についてゲーム理論モデルが適用可能であることまでは期待できても、そのモデルに具体的にどのようなパラメータを入れなければならないのかについては、まだほとんどわかっていな

いのだ。もちろん接触の結果起こりうるのは衝突、紛争だけではなく平和的共存と互恵主義の可能性についても理論は教えてくれるが、紛争かあるいは共存か、の分岐がどこにあるのか、を知るためには、こうしたパラメータについて知らなければならないはずである<sup>(13)</sup>。

#### 4. Existential Risk, Superintelligence, 宇宙植民

ところでこのような意味での「パラドックス」は、何もクーパーの「コンタクト・パラドックス」に限られるわけではない。というよりこれは先端科学技術一般のはらむリスクや、地球環境問題、大規模自然災害のリスクなどを含めた、長期的な人類の未来について考えたときに浮上してくる問題一般に通じる、普遍的な構造に従っていると言える。

ニック・ポストロムが提唱するExistential Risk, およびそのサブカテゴリとしてのGlobal Catastrophic Riskの概念は、まさにその構造を的確にえぐり出している<sup>(14)</sup>。ポストロムのいうExistential Riskは必ずしも急性の劇的な変化による人類絶滅に限定されるわけではなく、人類文明の不可逆的な衰退の緩慢な過程といった事象をも含むが、ポストロムがその認識論的な特徴として挙げるのは、Existential Riskが非常に頻度の低い、稀な事象であって、なおかついったんそれが発生すれば人類すべての絶滅や不可逆的衰退といった、取り返しのつかない結果をもたらすため、よくあるリスク、つまり高頻度で反復され、その損害も小規模で限定的であるような事象と異なり、それへの対策のノウハウを、経験を通じた試行錯誤によって学ぶことができない、という点だ。それに対しての準備は、あくまでも科学的な根拠を持った理論的予測に則るしかなく、またそれを予防するにせよ耐え忍ぶにせよ、失敗は許されない。

よく知られているようにポストロムはExistential Riskの例として、わかりやすい急性の破局的事象、つまりGlobal Catastrophic Riskに限っても、人為

的な事象としての核戦争、バイオテロ、バイオテクノロジーやナノテクノロジーの暴走、暴走温暖化や、あるいは非人為的な、純粋な自然災害としての天体衝突、超大規模噴火、あるいは（人為的なバイオテロではない自然発生的な）パンデミックなどを数え上げている。更にポストロムはテクノロジー暴走のヴァリエーションとして、人類の制御を離れて自律的に作動し、かつ自己修正による成長によって、そもそも人類による制御を不要にするだけでなく不可能にしてしまうAI機械、Superintelligenceの出現を挙げる。ポストロムによれば自律的な超AI機械の発展は、仮にそれが人類の急激な絶滅をもたらさないとしても、長期的にみれば人類の生存の余地を狭め、文明における人類のリーダーシップを掘り崩す可能性があるため、れっきとしたExistential Riskを構成するのである<sup>(15)</sup>。

そしてここまでの考察を踏まえるならば、クーパーが「コンタクト・パラドックス」の概念とともに示したような、ETI, ETCとのコンタクトもまた、れっきとしたExistential Riskを構成することは言うまでもない。

また宇宙倫理学の領域においては、ETI, ETCとのコンタクトを度外視し、人類（とその末裔）の宇宙進出に射程を限ったとしても、（擬似）「パラドックス」の例を見つけることができる。Space ColonizationないしSpace Settlement、つまり地球外、深宇宙における人類の恒久的生活拠点の大規模な開発の可能性を考えるならば、仮にETI, ETCの問題を無視したとしても（仮にコンタクト可能な時空内には我々以外の知的生命・文明は存在していなかったとしても）、ある種パラドキシカルな問題が浮上するのである。

稲葉振一郎は、リベラルな社会の枠組み、すなわち、私有財産制度と自由な市場経済を軸とした経済体制と、言論・思想の自由に立脚した自由民主主義が維持されるならば、環境の悪化等によって地球上での人類の生存が著しく困難になったりしない限りは（すなわちExistential Riskの回避のために不可避で

あるということにでもならなければ), 大々的な宇宙植民, 地球外の人類の恒久的生活拠点の大規模な発展が起こるとは考えにくい, と論じている<sup>(16)</sup>。すなわち, 誰にも強制されない市民の自発的活動によっては, あるいはたとえ強制的な統治権力によるとしても, あくまでも合法的な政治プロセスの結果実施される公共事業によっては, 無視し難い数の人々が自発的に地球の外に生活拠点を移し, そこに新たなコミュニティを築くようになるとは思えない, と。学術研究や資源獲得のための拠点は, あくまでも現代の極地基地や海底油田のリグのような, あるいはそれこそISSのような前線基地にとどまり, 大学都市や鉱山町のような独立して持続可能なコミュニティ, 都市, あるいは国家にはなりえないだろう, と。その主たる理由はやはり, 微小重力や宇宙線などを含めた宇宙空間の過酷さであり, そこから人間を保護できる人工環境を大規模に建設し維持するコストが耐え難いほど高くなるだろう, ということである(我々は人間をその中で問題なく生存させられるような, 自給自足の閉鎖生態系の管理技術さえまだ確立できていない)。

更に言えばこのような地球外の拠点で生活することは, たとえ太陽系内に限定したところで, インターネット以降の常時接続の高密度双方向同時通信によって繋がれた社会から己を切り離すこと, その利便性の大半を捨てることでさえある。人類史を省みるならば, 協力と取引の利益という, 集住を促す求心力と, 生業を営むための, また外敵から身を守るためのテリトリーの確保と分散を促す遠心力のバランスによって社会の空間的配置が決定されてきたが, 産業革命以降は明らかに, 都市化を進める求心力の方が優越している。20世紀末以降の情報通信革命は, 一見この傾向に歯止めをかけるかのように見えたが, 実際には必ずしもそうはなっていない。更にここに深宇宙の距離がもたらす通信遅滞・減衰のことを考えれば, 仮に地球外への移民が大々的に進んだとしても, 人口の大半は通信遅滞が秒単位で済む地球近辺(せいぜい地球一月系内)にとどまる可能性は大きい, と考えねばならないだろう。

仮にこれらの困難を勘案してもなお、宇宙への大々的な移住を有意味な選択肢として浮上させるような状況があるとすれば、ひとつには先述の、地球における人類の存続が困難ないし不可能となった場合であるが、他に考えられるのは、地球においてリベラルな社会が崩壊して全体主義化した場合、その抑圧を逃れて自由を守るために、という可能性である<sup>(17)</sup>。高速度高密度のデータ通信社会が、逃れがたい監視と統制の全体主義体制をもたらしたとしても、その腕は通信の遅滞／減衰によって、深宇宙のコロニーにまでは伸びることはないであろう。

しかし「自由で平和で安全な地球社会」という大前提が保たれている場合には、少数の冒険者であればともかく、大量の普通の人々が自発的に、過酷な宇宙に不可逆的に進出することは考えにくい。稲葉（2016）では、過酷な地球外環境に適応できるように自己改造したエンハンスドヒューマン、あるいは自立型AI機械（人造人間）であれば、リベラルな社会のもとでも自発的に宇宙に進出する可能性がある、と推論したが、そもそも外宇宙での活動以外に、そのような改造人間・人造人間への社会経済的需要があるとは考えにくい。すなわち、大規模な宇宙植民活動が行われるためには、まえもって大量の改造人間・人造人間が、一部の特別な専門家やエリートではなく、社会内で一定の層をなす（マイノリティではあっても）普通の人々として存在していることが望ましいが、大量の改造人間・人造人間が（リベラルな社会のもとで、強制によるのではなく自発的な選択の結果として）存在しているためには、そのような存在への需要が必要であり、大規模な宇宙植民活動以外にそのような需要を考えることが難しい、と。これは「パラドックス」とは言わないまでも、一種の悪循環ではないだろうか。

またSF作家小川一水は、大河小説『天冥の標』の中で「移民団のパラドックス」を提示している<sup>(18)</sup>。すなわち、太陽系外、恒星間宇宙への大々的な進出、宇宙植民は、世代間宇宙船を使うにせよ、あるいは生体を冬眠させるにせよ、



またあるいは種子やあるいはデジタルデータ化して保存するにせよ、不確実な成果のために膨大なコスト、とりわけ膨大な時間を必要とする。外界から完全に孤立した世代間宇宙船の閉鎖社会の中で生きる場合にはもちろんのこと、人工冬眠やデータ化によって、移動中、宇宙航行中の主観時間をゼロにした場合にも、その間に故郷を含めた通常空間で使えたはずの膨大な時間、その結果生じる故郷とのタイムラグといったコストを払わねばならない。そうしたコストを支払ってでも故郷を離れて宇宙に進出する動機としては、場合によってはそもそもその存在の有無自体が不確かな、目的地で期待される成果だけでは不十分なのではないか、と小川は問う。そこで小川は、このような宇宙植民が実行されるためには、目的地への到達への期待だけではなく、過酷な宇宙航行それ自体、つまり目的地への到達ではなくその途中の過程自体が、自己目的化されなければならないのではないかと論じる。しかし仮にこのような航行それ自体の自己目的化が完成してしまえば、本来の目的の達成、つまり目的地への到達自体の意味が見失われてしまいかねないだろう。たとえば世代間宇宙船を安定的に航行させ、閉鎖社会をうまく維持することが可能になれば、そこから降りる動機の大半が失われてしまうのではないかと。「遠方への困難な移民に挑戦しようとするほど、外部と隔絶した閉鎖生活を独力で営む仕組みができあがり、移民する意味が失われてしまうという矛盾。」(小川(2011))これはいくつかの古典的なSF作品で問われた問題である<sup>(19)</sup>。

## 5. 分岐する人類

稲葉(2016)が提示した宇宙植民と改造人間・人造人間の悪循環と、小川の「移民団のパラドックス」はいずれも「フェルミ・パラドックス」への解答候補として——知的生命の本格的な宇宙進出が行われる可能性は低い、という主張として解釈することができるのはもちろんであるが、ここではその点につい



て掘り下げることはしない。またそのような議論である以上、それらを主體的に真面目に論じることは、「コンタクト・パラドックス」それ自体の追究とは両立しない。そうではなく、ここで問題としたいのは、これらの議論とクーパーの「コンタクト・パラドックス」、そしてポストロムが論じている、スーパーインテリジェンスの出現を含めたExistential Riskの認識論の間に、共通する（擬似）パラドックス的な構造がある、ということだ。つまりは「その課題を解決するための必要条件が、その課題の達成を通してではないと満たされないのではないかと疑われる」、という構造が。

これは単なる偶然ではない。我々の考えるところでは、こうした構造が浮上する重要な理由のひとつは、おそらくは異質な存在としてのETI（や人工知能）に対峙する「人類」自体の同一性が、十分に疑われないままにそこで自明のものとして前提とされているからである。これらの議論においてはいずれも、そこに提示されている困難な問題に立ち向かう主体としての人間、人類の同一性が、自明なものとして疑われないままに前提とされているのではないのか？ 逆にそのことによって、ETIにせよAIにせよ、人類にとっての困難な課題として想定されるものの、それらの人間にとっての異質性、他者性が際立たされ、また反対に人間の同一性、不変性が強調されて、人間とETI、AIとの対峙の困難さがいたずらに強調されがちになるのではないのか？ つまり、そこで想定される他者性とはまさに、その定義上、同語反復的に「人類とは異質で理解し難く手に負えないもの」とされがちだということだ。ETIに対峙する、地球出自の知的生命の一種としての人類、あるいは人工知能に対する、自然発生的な生物学的知能の担い手としての人類の自己同一性が、あまりにも自明視されているからこそ、逆にそれとは異質ななにかとの対峙が困難としてイメージされてしまうのではないのか？

だが実際には人類は決して一枚岩ではない——過去においてもそうだった

し、おそらくは未来においてはなお一層そうだろう。

宇宙植民、ETIとのコンタクト、スーパーインテリジェンス、いずれも未決の未来における、実現するかどうかさえ定かではない可能性である。それらの課題に直面する我々の末裔のことを、我々と同一の人類、我々と本質的に同じものの単なる延長と考えるから、これらの難問が擬似パラドックスとして現在の我々の眼には映るのではないだろうか。しかしながら、これら諸問題の多くが現実的に浮上する頃には、今現在生きている我々はおそらく誰一人として生きてはいない。それらの問題に直面する具体的な「誰か」は、我々の生物学の子孫であるかもしれないし、仮にそうではなくても、文化や知識の継承者ではあるだろう。しかし現に生きている我々自身ではないはずだ（トランスヒューマニストは個体としての不死を獲得することによって、この想定を破ろうとするわけだが<sup>(20)</sup>）。だとすればそうした人類の末裔、後継者たちもまた、いま現在の我々にとっては何程か異質な他者であり、そうするとそれらとETIやスーパーインテリジェンスとの違いは、程度問題であると言える（それは今世紀中にトランスヒューマニストが期待するような不死が実現した場合についてさえ言えるかもしれない。数百年、数千年を個体として生き延びた人は、果たしてどの程度我々の知るヒトと同じだろうか）。そもそも人類が作るスーパーインテリジェンスでさえ、仮にそれが生物学的人类を滅ぼすことになったとしても、今現在の我々人類の文化や知識の継承者でなかったら何だというのか？ このように、現代の我々人類と、未来のその継承者との間の同一性が自明ではないものとされるならば、逆にそれとETI、ETCとの差異もまた、それほど隔絶したものと考えなければならぬ理由はないのではなかろうか？

稲葉（2016）の議論を宇宙植民への懐疑論としてではなく、その可能性について肯定的に考える論法として解釈するならば、それは遠い将来（少なくとも千年、一万年単位の未来）において、我々人類の末裔が、その生物学の子孫はもちろん、ある種の自律型AI機械などを含めて、互いに異質ないくつもの集

団へと分岐していくことを予想する。そのような未来の人類の末裔は、今現在の我々と全く異質ではないにせよ、相当程度異なった存在であるだろうし、またそれらの中でも互いに他者同士であらざるを得ないだろう。そしてこのような可能性を許容すること、ないしは促進することは、実は「コンタクト・パラドックス」やその他のExistential Riskの困難をある程度緩和することになるのではないだろうか。

なぜならば、この場合、ETIとの異質性に悩む前に、人類の末裔は自分たちの中での異質性と対立に悩まねばならない。こうした問題への対処は同時に、地球出自ではないETI、ETCとの本格的なコンタクトを前にした予行演習となりうるだろう。またそうした異質な多数の集団に人類の末裔が分化してしまえば当然、他の種類のExistential Riskもまた相対化される。すなわち、異質なグループに分化した人類の末裔たちがそれでも互いに完全に疎遠にはならず、ゆるやかにではあれ知識共有を続けるならば、一惑星、特定の環境に縛られているがゆえにやり直しがきかないはずのExistential Riskに対して、試行錯誤が可能となる<sup>(21)</sup>。

もちろん以上の議論は、そもそも大げさであるとも言える。先にも論じたように、ETI、ETCもこの宇宙に存在するものである以上、その法則に拘束されているはずだ。だからそれらは、たとえ我々の知らない数学を知っていたとしても、我々が証明した定理を覆したりはしないはずだ。彼らの作る計算機も、というより彼らの知性もまた、我々のそれと同様に万能チューリング機械の一例だろうから、それにできることは何でも（しかも我々より上手に）できるかもしれないが、それにできないことは決してできない。またそれらは我々の知る物理法則を破ることもないだろう——もし彼らの技術が我々にそのように見えたとしても、それはただ彼らが、我々がまだ知らない、より根底的な物理法則を知っていて、そのより根本的な法則によれば、我々の知る範囲の外側で我々

がまだ知らない事が起こり得て、それが我々の既知の物理法則では説明がまだつかない、というだけのことである。だとしたらETI, ETCがどのような存在か、何をなしうるか、を推測する際に、我々の既知の数学や物理学、あるいは生物学や経済学、哲学の知見が役に立たないはずはない。それは決して十分な理解には到達できないにしても、決して無駄ではない。

また、産業革命以降の様々な技術革新、その背後にあるより基礎的な学術研究、そして個人の個性の自由な発露として独創性を競うようになった近代以降の芸術制作は、基本的には「誰にも必要とされなかった、どこにも需要がなかったはずのものを新たに作り出し、あとからその需要を引きずり出す」こと、「その到来の必要条件が満たされていなかったはずのものを作ってあとからその存続のために必要条件をでっちあげる」ことを中心としていたのではないか？ そのように考えるならば、先程は「結局は誰もそれを特別欲しそうにないので、リベラルな社会のもとでは起こりそうにない」と論じた宇宙植民さえも、大々的なビジネスとして展開する可能性は排除できない。また、このようにしばしば意図せずして新たなものを作ってしまうことが近代的な意味での「創造」であったことを思えば、それもまた、ETI, ETCとのコンタクトとは幾分違った意味においてではあれ、「未知との遭遇」であるには違いない。

このような論法は、デイヴィッド・ドイッチュがポストロムのスーパーインテリジェンス論や類似的AI脅威論、シンギュラリティ論に対して「どんな超AIも万能チューリング機械の一種でしかない」と呈した疑問とも通じている<sup>(22)</sup>。実際ドイッチュが主張するように、「AIだろうと自然発生的な知的生命だろうと、どのような知性の営みも万能チューリング機械の一種でしかない」である以上、そこに絶対的な異質性、理解不能性というものはおそらくは存在しない。ただしその意味で根本的に質的な断絶は超AIやETI, ETCと現在の我々との間に存在しないとしても、あまりに量的に巨大な差——それを埋めることは不可能ではないが、膨大な労力が必要とされるような差——というもの

は存在しうる。それはたとえば昆虫とヒトの知性の間にあるだろう差か、それよりも大きなものかもしれない。ポストロムが際限なく指数的に成長する超AIと私たち人類の自然知能との間に見出したのもそのような、いやそれ以上の差である。ただ一方我々は、明らかに知性の上で我々より劣っているだろう昆虫を、それでも完全に理解し好き勝手に支配できているわけではもちろんないのだ。同様の可能性を、人類と超AI, ETI, ETCとの関係に期待することは不合理ではないだろう。

### \* 補論

非常に単純に言えば、3. で提示した立場bを元に「コンタクト関数」 $a$ とでもいうべきものを考えることもできるだろう。bそれ自体に忠実に考えるなら「ETI発見関数」となるだろうが。つまりある労力、コストを投入すればそれだけETIの発見確率なり、ETIとのコンタクト成功率が上がる、というものだ。そしてこれとは別に、コンタクトそれ自体がもたらす損益の関数というものも考えられる。ここではそれを、損益を独立変数とし、ある一定の損害／利益を人類社会にもたらすコンタクト事象の確率分布というものを考えるのだ。もちろんコンタクト関数（ETI発見関数）の方からも、コンタクト（発見）成功に至るコストの確率分布が得られる。その上で単純化して、この二つの確率分布は互いに独立である、と考える。そうすると、SETIに投入されたある一定の労力がもたらしうる、成功したコンタクトから発生する損益の確率分布は、単純に両者を掛け合わせる形で得られるだろう。

ものすごく単純化してコンタクト関数 $P_c(r)$ というものを考える。これは、SETIにコスト $r$ を投入したときのコンタクト成功率を表す、と考える。 $r$ を何で測るかは後述する。この関数がどのような形状をしているか、については、 $r$ の単調増加関数とし、かつ増加率は逡減する、と考える。数学的に表現すれば、

$\frac{dPc}{dr} > 0, \frac{d^2 Pc}{dr^2} < 0$  である。従属変数が確率変数であるから当然に  $0 \leq Pc(r) \leq 1$  で

あり、もし何もしていなくとも、ETIの方から地球人類を発見し、コンタクトしてくる確率が0ではないとするなら、その確率を $q$ とするなら、 $q \leq Pc(r) \leq 1$ となる。

これに対してコンタクト事象の損益の分布  $\beta$  を考えるのは難しいが、期待しうる利益については上限がない、無限大である、と考え、他方起こりうる損害については下限というものがある、と考える。ここでの下限とは例えば人類絶滅である。人類が絶滅しても地球の生態系は保存される、といった場合と、人類ごと地球も消滅する場合とでは、考えようによっては優劣をつけることに十分に意味があるが、ここでは単純に人類絶滅のみをワーストケースとする。更に単純に考えるために、損害と利益を同じ尺度で測り、損害をマイナスの利益、と考えることにする。その上で起こりうる最大の損害、つまり人類の絶滅を、非常に単純に、人類が動員できる資源すべて（そこには他ならぬ人間自身の知識、能力、生命身体もはいる）の喪失、と考え、かつそれを乱暴に一次元的な数値で表現できる、とする。これはある種の功利主義の理論枠組みを援用したものである、と考えていただいてもよい。この人類が動員できる資源総量を数値化したものを  $R$  とする。とすると考えられる最悪の損害とは  $-R$  であり、コンタクトの損益の分布関数は、半开区間  $[-R, \infty)$  上で定義される。その上で、先にコンタクトに投入される労力、コストを  $r$  で表現したが、これは当然この  $R$  と単位、次元を同じくするわけであり、 $0 \leq r \leq R$  と考えることが理にかなっている。だからこの分布関数を  $Pb(r)$  と表現しても構わないようにも思われるが、概念的には、コンタクトに投入される資源総量と、コンタクトによって失われる資源総量とは当然に別個のものなので、別の記号で表現したほうが無難であり、とりあえず  $Pb(b)$  としよう。つまり  $-R \leq b < \infty$  である。この  $Pb(b)$  が具体的にどのような形をしているか、想像することは難しいが、ここでも非常に

単純化して、最悪の事態になる可能性はそれほど高くはなく、かといって良い成果が得られる可能性の方もどんどん低くなる、という平凡な形状を考える。この分布の左側の裾野は短い、右側の裾野は長い。具体的には対数正規分布に似た形をイメージすればよいだろう。ただしここで重要なのは形状だけではなく、その分布の平均値や中央値がどのあたりになるか、だろう。

ここでもっとも単純にリスク中立的な意思決定原理に従うならば、人類は損益 $b$ が最大化されるようにSETIなりMETIなりを行うべきということになるが、更に言えばごく特殊な場合、たとえば何もしなくてもETIによってコンタクトされてしまう確率が無視し難いほど大きい場合などを除けば、損益 $b$ の期待値がマイナスになるようなことはしてはならない、といえる。ここで期待損益を $E(b)$ とすると、 $E(b) = \int_{-R}^{\infty} bPb(b)db$ と表現できる。これとは別に損益の

中央値 $Me(b)$ や最頻値 $Mo(b)$ についても考えることができる。損益分布が上述のように対数正規分布で近似できるような形をしていた場合、 $E(b) > Me(b) > Mo(b)$ となる。リスク中立的ではなく、リスク回避的な——つまり利益は少なくなっても、人類滅亡を含めた危険の回避を重視する意思決定原理を採用する場合には、期待利益を最大化するよりも、中央値や最頻値の方がベンチマークとしては適切だということになる。付言すると、中央値は、それ以上の利益を得られる確率と、それ未満の利益しか得られない確率とがちょうど50%ずつで均衡するポイントである。また最頻値は確率密度が最大となる、つまり最も実現可能性が高い状態に対応すると解釈することもできる。

繰り返すが、コンタクトの成功確率と、コンタクトが成功した場合の損益の確率分布とは互いに無関係で独立だと想定すると、仮にリスク中立的にコンタクトの利益を最大化することを目指すとするれば、これは

$$\begin{aligned} & \text{Max } E(b)Pc(r) - r \\ & \text{s. t. } 0 \leq r \leq R, \end{aligned}$$



つまり最大利用可能な資源の上限が $R$ という制約のもとで、コンタクトの期待純利益、つまりコンタクトの期待粗利益と投入されたコストの差を最大化する、という風に解釈できる。ただしここで注意すべきは、そもそも期待損益がマイナスである場合には、何もしないことが期待利益の最大化、つまり期待損害の最小化になるということである。これはベンチマークに損益の期待値の代わりに中央値や最頻値を用いた場合にも同様である。先に見たようにコンタクトの損益分布が想定したようなものであれば、期待利益がプラスであっても、損益の中央値や最頻値がマイナスという場合も考えられる。その場合は上述のようなリスク回避的な意思決定基準のもとでは、たとえその期待利益がプラスであっても、コンタクトはそもそも目指されるべきではない、という結論が出る。それだけではない。仮に中央値や最頻値をベンチマークとした基準のもとでの理論的な利益がプラスになったとしても、そこから導き出されるコンタクトへの最適な努力水準は、リスク中立的な期待利益志向の場合のそれに比べて低くなるはずである。

標準対数正規分布  $f(x) = \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{(\ln x)^2}{2}\right)$  のグラフを以下に提示する

(図1)。ここで期待値(平均値)  $E(b)$  は  $e^{0.5} \cong 1.648721271$ 、中央値  $Me(b)$  は  $e^0 = 1$ 、最頻値  $Mo(b)$  は  $e^{-1} \cong 0.367879441$  である。

まずはこのグラフを素直に、上述の、コンタクトの損益の確率分布を概念化

したものと解釈しよう。つまり  $Pb(b) = \frac{1}{\sqrt{2\pi b}} \exp\left(-\frac{(\ln b)^2}{2}\right)$  である。ここで

は単純に  $R=0$ 、つまりコンタクトそれ自体の損害は0と想定するところから始めよう。プラスの損害がある場合には、このグラフを左側(マイナス側)に

最大損害  $R$  の分だけシフトすればよい  $Pb(b) = \frac{1}{\sqrt{2\pi}(b+R)} \exp\left(-\frac{(\ln(b+R))^2}{2}\right)$

とする。



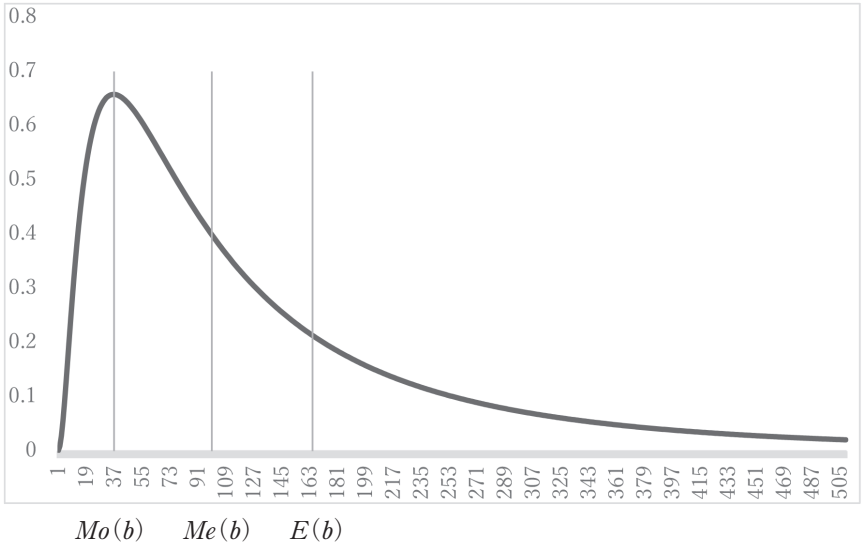


図 1

次にコンタクト関数 $Pc(r)$ についても、イメージしやすいように特定しよう。

単純化のために $q=0$ とし、 $Pc(r)=\frac{r}{1+r}$ とすれば、 $r \geq 0$ の範囲で、 $Pc(0)=0$ ,

$\lim_{r \rightarrow \infty} Pc(r)=1$ となり、更に $\frac{dPc}{dr} = \frac{1}{(1+r)^2} > 0$ ,  $\frac{d^2 Pc}{dr^2} = -\frac{2}{(1+r)^3} < 0$ であるので先

の想定を満たす。

リスク中立的に期待利益の最大化を目指すとする、制御できる変数は $r$ のみなので、 $E(b)$ は定数として扱えて、最大化の条件は

$$E(b) \frac{dPc}{dr} - 1 = 0$$

なので、

$$E(b) = (1+r)^2$$

となる。

「コンタクト・パラドックス」とその同類たち

そうすると、期待利益を最大化する投入コストを $r_e$ とすると、

$$r_e = \sqrt{E(b) - 1} = e^{0.25} - 1 \cong 0.284025417 > 0$$

となり、有意な解が現れる。しかしよりリスク回避的な決定基準をベンチマークとして、可能な利益の平均、期待値の代わりに、可能な利益の中央値 $Me(b)$ を目指したとしよう。そうすると、

$$r_e = \sqrt{Me(b) - 1} = e^0 - 1 = 0$$

で、ここでの設定だとたまたま0となり、コンタクト努力をしようがすまいがどちらでもよい、という結論となってしまふ。更により慎重に、より強くリスク回避的な意思決定原理として、最頻値 $Mo(b)$ を目指す場合には、その値自体は正であるにも関わらず、最適投入コストは

$$r_e = \sqrt{Mo(b) - 1} = e^{-1} - 1 \cong -0.632120559 < 0$$

で負となってしまう、コンタクト努力は行わないほうがよい——コンタクトの成果はその労力に見合わない可能性が高い、という結論となる。これは当然、 $R > 0$ の場合を含めて損益分布のピークがシフトする場合や、コンタクト関数の形状の違いによって条件は変化しうるが、大体の方向性は明らかであろう。

\*註

- (1) ex. Cooper (2019=2021), Green (2022)。
- (2) いわゆる「フェルミ・パラドックス」をめぐる論争。ex. Webb (2015=2018)。
- (3) 国際政治学の抑止理論の星間戦争への敷衍を試みるKorhonen (2013) はこれに加えて、銀河系内の恒星の年齢分布から考えると、我々の太陽系は比較的若い方に属するがゆえに、我々人類もまた銀河系内の知的生命・文明の分布の中では若い方に属する可能性が高い、と指摘する。
- (4) 著名な進化生物学者のDiamond (1991=2022) は人類史からの類推で強い悲観論をとる。Baum et al. (2011) はシナリオ分析によって、ETI, ETCとの接触が人類にとつ

## 「コンタクト・パラドックス」とその同類たち

て有益である可能性、有害である可能性双方について整理し概観しているが、敵対的ではない、協調的ETIとの接触においても、意図せざる危険が無視できないことに触れており、その中には感染症リスクやコンピュータ・ウィルスの他に文化接触のショックによる人類の道徳的腐敗・墮落の可能性も数え入れられている。

- (5) 代表的な楽観論者としてAlbert A. Harrison, Douglas A. Vakochが挙げられる。Vakoch and Harrison eds. (2011) 他。
- (6) Bostrom (2002), 他。
- (7) 例外として, Jebari and Olsson-Yaouzis (2018) は, 特定のETI, ETCの存在を知ること自体が, そこへの信号の送信を可能にしてしまい, それが可能である以上, 何らかの事故や悪意, あるいは軽率さによって実行されることがありうる, と論じており, これはMETIのみならず受動的な通常のSETIへの慎重論としても解釈しうる。
- (8) Jebari and Olsson-Yaouzis (2018) がその例外である。
- (9) Green (2021) はアメリカ合衆国の世論調査の「おおむね人口の過半数がETIの存在を信じている」との結果を, またPeters (2011) における「複数の宗教にまたがる回答者の8割が, ETIの発見が自分の信仰に影響をおよぼすことはない, という結果を紹介している。
- (10) Jebari and Olsson-Yaouzis (2018) の立場がこれである。
- (11) これについては, 国際政治学における「デモクラティック・ピース」論 (Russett (1993 = 1996) 他) をもとに 基本アイデアを提示したHarrison (2000) 以降, いくつかの論文が書かれている。Korhonen (2013) はHarrisonほどの楽観論はとらないまでも, 文明間交流における交易などの互恵性の成立の可能性を指摘した上で, 国際政治学の抑止理論をもとに, 星間戦争の可能性は否定しないが, 攻撃自体の困難性に加えて報復の可能性が, 恒星間戦争においても抑止メカニズムを成り立たせうる, と論じる。またKorhonenは, 恒星間航行が可能な技術レベルの文明においては, (たとえば恒星間航行可能な有人宇宙船は完全な資源リサイクルに立脚しているはずであるから) 高い資源利用効率が達成されているはずであるため, 物質的な資源獲得などを旨とした貪欲な征服指向は星間戦争の動機となるとは考えにくく, なるとしたら侵略への恐怖からくるパラノイアックな先制第一撃指向であろう, と推論する。

これに対してJebari and Olsson-Yaouzis (2018) は, Korhonenと同様に星間戦争の主要動機を侵略に対する恐怖, 生存への渴望に求めるが, W・V・O・クワインやドナルド・デイヴィッドソンの言語哲学, とりわけ「根源的翻訳・解釈」理論 (Davidson (1973=1991), Quine (1964=1990)) をもとに, 同一の生息環境を共有しない異種知性同士の間では, 翻訳による相互理解が実質的に不可能である, と論じ, 互いにまともなコミュニケーションが成立しないので抑止戦略は成り立たず, ゲーム理論的には先制第一撃が最適解となる, と論じる。ここからJebari とOlsson-YaouzisはMETIの

否定のみならず、受動的SETIに対してさえ消極的な態度を表明する。このような議論はSFを含めてもちろん珍しいものではなく、たとえば劉慈欣『三体』三部作第二部『黑暗森林』（劉（2008=2020））の基本アイデアである。Webb（2015=2018）においても「フェルミ・パラドックス」への回答としてこのように隠れ住む文明の可能性が検討され、たとえば高度文明にとってブラックホールはエネルギー資源としても隠れ家としても好適である、と指摘されている。

- (12) 行動経済学についてはKahneman（2011=2012）他。
- (13) このような問題意識を踏まえてDöbler and Raab（2021）は、SETIを射程に入れた普遍心理学としてのexopsychologyの樹立を提唱する。
- (14) Bostrom（2002）。
- (15) Bostrom（2014=2017）。
- (16) 稲葉（2016）。
- (17) 稲葉（2019）。
- (18) 小川（2011）。
- (19) ちなみに小川自身はこのパラドックスに対して、どのような答えを与えているのか？ 『天冥の標』の世界においては、宇宙は知的生命、星間文明で溢れかえっており、恒星間宇宙、いや銀河団のレベルで熾烈な生存闘争が続いているため、そもそもこの隘路を突破して星間文明を樹立しない知的生命、技術文明は長期的には存続できない。そこでは「フェルミ・パラドックス」は人類がまだ若い知性種族であるがゆえの錯覚、泡沫の夢にすぎない。それゆえ「移民団のパラドックス」の含意は、『天冥の標』という作品の中では深く掘り下げられることはない。
- (20) ex. Kurzweil（2005=2007）。
- (21) Webb（2015=2018）においても包括的にサーベイされているとおり、人類が近い将来にETI、ETCと遭遇する可能性は低いと推測される。具体的に定量的評価を試みた最新の研究であるSong and Gao（2022）においては、銀河系内に範囲を限定して、楽観的推定（地球型惑星に知的生命が文明を建設する確率が0.1%、文明形成時における恒星の年齢がその寿命の25%）において、人類がETIからの最初の通信を受信するまでの期待時間は2,000年程度、悲観的推定（地球型惑星に知的生命が文明を建設する確率が0.001%、文明形成時における恒星の年齢がその寿命の75%）においては40万年程度、とされている。悲観的推定に立った場合はもちろん、その中間のオーダーとして1万年単位の時間を見込んだ場合にも、もし仮に人類文明がそこまで生存していたとしたら、既に恒星間宇宙に広がり、複数のコミュニティに分岐している可能性が高い、と考えられる。

これに対してKaku（2018=2019）は「DNA時計」を踏まえて、たとえ互いに交流のない集団に分岐したとしても、それらの間でのDNAの相違は10万年に0.1%程度に

しかならない、と推定し、仮に恒星間宇宙に広がったとしても人類の間での異質な集団への分岐は起こりにくい、と論じる。ただしそこでは人為的介入の可能性は考慮されていない。カクは人類が自発的に自己改造して異質な存在に変貌すること、互いに異質な集団へと進んで分岐していく可能性は低い、と見積もっている。カクは「穴居人の原理」と呼んでいるが、要するに「性選択」のメカニズムによって人は互いに好感を抱きうる範囲の姿かたちを保ち続けるだろう、ということだ。しかしまさに「性選択」、つまり周囲の自然環境ではなく異性の個体、つまりは同種のコミュニティ、社会環境に適応的な形質が選択されるというメカニズムこそがflamboyant displayとしての文化、芸術の進化的起源である、という理解はDiamond (1991=2022)を含めて広く流布している。稲葉 (2016) では実用的な観点から大規模な人体改造が普及する可能性は低い、と論じたが、近未来における有人宇宙飛行の大口需要として見込まれるのが、宇宙飛行それ自体を自己目的化した観光であると同様に、自己目的化されたファッションとしての人体改造の普及の可能性は否定できない、と考える。

(22) Deutsch (2011=2013)。

#### \* 参考文献

- Baum, S. D. Haqq-Misra, J. D. and Domagal-Goldman, S. D. (2011) "Would Contact with Extraterrestrials Benefit or Harm Humanity? A Scenario Analysis." *Acta Astronautica*, 68.
- Bostrom, Nick. (2002) "Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards." *Journal of Evolution and Technology*, 9.
- Bostrom, Nick. (2014) *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press. = (2017) 倉骨彰訳『スーパーインテリジェンス』日本経済新聞出版社。
- Cooper, Keith. (2019) *The Contact Paradox*. Bloomsbury Publishing. = (2021) 斎藤隆央訳『彼らはどこにいるのか』河出書房新社。
- Davidson, Donald. (1973). "Radical interpretation." *Dialectica*, 27. = (1991) 金子洋之訳「根源的解釈」。ドナルド・デイヴィッドソン『真理と解釈』勁草書房、所収。
- Deutsch, David. (2011) *The Beginning of Infinity*. Penguin. = (2013) 熊谷玲美・田沢恭子・松井信彦訳『無限の始まり』インターシフト。
- Diamond, Jared. (1991) *The Third Chimpanzee*. Hutchinson Radius. = (2022) 長谷川寿一・長谷川眞理子訳『第三のチンパンジー』日本経済新聞出版社。
- Döbler, Niklas A. and Raab, Marius. (2021) "Thinking ET: A discussion of exopsychology." *Acta Astronautica*, 189.
- Green, Brian Patrick. (2022) *Space Ethics*. Rowman & Littlefield.

- Harrison Albert A. (2000) “The relative stability of belligerent and peaceful societies : implications for SETI.” *Acta Astronautica*, 46.
- 稲葉振一郎 (2016) 『宇宙倫理学入門 人工知能はスペース・コロニーの夢を見るか?』ナカニシヤ出版。
- 稲葉振一郎 (2019) 『銀河帝国は必要か? ロボットと人類の未来』筑摩書房。
- Jebari, Karim and Olsson-Yaouzis, Niklas. (2018) “A Game of Stars : Active SETI, radical translation and the Hobbesian trap.” *Futures*, 101.
- Kahneman, D. (2011) *Thinking Fast and Slow*. Farrar, Straus and Giroux. = (2012) 村井章子訳『ファスト&スロー』早川書房。
- Kaku, Michio. (2018) *The Future of Humanity: Terraforming Mars, Interstellar Travel, Immortality, and Our Destiny Beyond Earth*. Doubleday. = (2019) 斎藤隆央訳『人類、宇宙に住む』NHK出版。
- Korhonen, J. M. (2013). “MAD with aliens ? Interstellar deterrence and its implications.” *Acta Astronautica*, 86.
- Kurzweil, Ray. (2005) *The Singularity Is Near: When Humans Transcend Biology*. Viking. = (2007) 井上健・小野木明恵・野中香方子・福田実訳『ポスト・ヒューマン誕生』NHK出版。
- 劉慈欣 (2008) 『三体Ⅱ 黑暗森林』重慶出版社。 = (2020) 大森望・立原透耶・上原かおり・泊功訳『三体Ⅱ 黑暗森林 (上・下)』早川書房。
- 小川一水 (2011) 『天冥の標Ⅴ 羊と猿と百獮の銀河』早川書房。
- Peters, Ted. (2011) “The Implications of the Discovery of Extra-Terrestrial Life for Religion,” *Philosophical Transactions of the Royal Society A* 369.
- Quine, W. V. O. (1964). *Word and object, studies in communication, first printing edition*. MIT Press. = (1990) 大出晃, 宮館恵訳『ことばと対象』勁草書房。
- Russett, B. (1993) *Grasping the Democratic Peace*. Princeton University Press. = (1996) 鴨武彦訳『パックス・デモクラティア』東京大学出版会。
- Song, Wenjie and Gao, He. (2022) “The Number of Possible CETIs within Our Galaxy and the Communication Probability among These CETIs.” *The Astrophysical Journal*, 928
- Vakoch, Douglas A. and Harrison Albert A. (eds.) (2011) *Civilizations Beyond Earth: Extraterrestrial Life and Society*. Berghahn Books.
- Webb, Stephen. (2015) *If the Universe Is Teeming with Aliens ... WHERE IS EVERYBODY? Seventy-Five Solutions to the Fermi Paradox and the Problem of Extraterrestrial Life 2nd ed.* Springer. = (2018) 松浦俊輔訳『広い宇宙に地球人しか見当たらない75の理由 フェルミのパラドックス』青土社。

「コンタクト・パラドックス」とその同類たち

\* 第一稿 2022年6月15日

\* 再修正 2022年11月18日

\* 本研究はJST, RISTEX, JPMJRX21J 1の支援を受けたものである。