

## 記憶喪失と世界喪失

### (Brains in a Vat and Vanishing of the World)

---水槽脳になったばかりの人が持つ記憶は元の世界を指示できるか?---

柴田正良 (金沢大学人文学類教授・哲学教室)

#### 1. 問題の所在

水槽脳 (Brains in a Vat) の仮説は、その根を辿れば、デカルトの夢の懐疑あたりまでは簡単に行き着く由緒正しき懐疑論の一つだ。ただ、デカルトの懐疑と異なる点は、これが現代の脳科学やコンピュータ・サイエンスの華々しい実績に裏打ちされ、いまや相当に現実味を帯びて実感されることである。水槽の中の脳という設定は、現代においては、脳という生物的認識装置がすべての認識の最終的な担い手であるという科学的な見方に根拠をおいている。それは、認識の生起を自然界の因果的メカニズムと捉えることの一つの極端な帰結だ。それによれば、あらゆる外界の刺激は脳の状態変化を引き起こすための原因にすぎず、原因は結果に似ている必要は少しもない。事実、赤いリンゴの知覚が処理される脳の神経細胞は赤くもならないし、リンゴの形にもならない。したがって、視覚原因が a と b という別々の刺激だろうと脳状態の結果が同じなら、認識は a と b を区別せずにそれらを例えば同じ c とみなすことだろう。かくして、現代の懐疑論は、脳の最終産物<認識された世界>がその原因である<実在の世界>と一致しない可能性を、誇大に描き出すところに成立する。つまりこの懐疑論は、大森荘蔵の言葉を借りれば、脳が<現実>や<実在>を生み出すという脳神話から生じるのだ (大森、[1981] 176)。本稿は、この水槽脳仮説の中でも、パトナムが提示した反懐疑論の論点のある局面だけを取り上げる(1)。つまり、全体としてパトナムの反懐疑論論証が成功しているか否か、あるいは、それは形而上学的実在論に対する反論としてどう評価されるべきか、といった点には触れない。

しかし、先を急ぐ前に、パトナムによる水槽脳仮説の設定を簡単に確認しておこう。パトナムはこう述べる。

ある人 (あなた自身と考えてもよい) が邪悪な科学者による手術を受けたと想像せよ。その人の脳 (あなたの脳) は身体からとりはずされ、脳を生かしておくための培養液のはいった水槽に入られている。神経の末端は超科学的コンピュータに接続され、そのコンピュータによって、脳のもちぬしはすべてがまったく平常通りだという幻想をもたされる。人々も、いろいろな対象も、大空や大地も、みなあるように思われる。しかし本当は、その人 (あなた) が経験していることはみな、コンピュータから神経末端に伝わる電子工学的インパルスの結果なのだ。そのコンピュータは非常に賢くて、あなたが手を上げようとすると、あなたは自分が手を上げているのを「見」たり、「感じ」たりすることになる。さらに、プログラムを変えるだけで、邪悪な科学者は彼の思い通りに、どんな状況や環境を

も、犠牲者に「経験」させる（幻覚を生じさせる）ことができる。脳手術の記憶を抹消することもできるから、犠牲者には、自分が座ってこんなお話を読んでるようにさえ思えるかもしれない。それは面白いけれどもまったく馬鹿げた想定で、邪悪な科学者がいて、彼は人々の脳を身体からとりはずし、脳を生かしておくための培養槽の水槽に入れるのだ。その想定によれば、神経の末端は超科学的コンピュータに接続されており、そのコンピュータによって、脳のもちぬしはすべてがまったく平常通りだという幻想を・・・(Putnam [1981], 5-6、邦訳、7-8 頁)。

パトナムは、以上の設定から、仮にある人が水槽脳となった場合、その人の言明「私は水槽脳である」は自己論駁的 (self-refuting) 言明となり、それゆえ水槽脳仮説は必然的に偽となると論ずる。ちなみに「自己論駁的」想定とは、それが真であることがそれ自身の偽であることを含意するものである」(ibid. 7、邦訳、10 頁)。しかも、さらに、この想定は、それを考えることも口に出して主張することもできないような種類のものとされる。仮にこのパトナムの論証が成功しているとすれば、この種の懐疑論は決定的に退けられたことになるだろう。もちろん、パトナムのややラフな形でのこの論証に対しては、いくつかの反論が提出されている。本稿は、それらの反論に対する直接の再反論を試みるものではないが、<指示の及ぶ範囲>を再考することによって、間接的なパトナム擁護論を意図するものである。

パトナムの論証の前提になっているのは、ある種の「指示の因果説」といってよいであろう。パトナムは定義の形で明確には述べていないが、それは、少なくとも、指示する項目（言語記号、絵、心的表象など）と指示される項目との間に何らかの因果的な結びつき (causal connection) が存在することを要請しているように思われる。両者の間に存在すべき因果的結びつきがどのようなものだと想定されているのかは明瞭ではないが、パトナムが論難する魔術的指示が、思考や意図や記号と事物との間に一切の因果的結びつきがないにもかかわらず可能となる魔術的結びつき (magical connection) に依るとされていることから、何らかの因果関係が指示関係成立の必要条件として考えられているのは間違いないであろう。例えばパトナムは、彼の議論の前提として、生まれてからずっと水槽脳であった人が行う指示に関して、こう述べているからである。「彼の思考は現実の樹木を指示しない。これから議論するいくつかの理論において、彼の思考はイメージの樹木を指示するかもしれないし、樹木経験を引き起こす電子工学的インパルスを示すかもしれないし、そのような電子工学的インパルスの原因であるプログラムの特質を示すかもしれない」(ibid. 14、邦訳、20 頁)。というのもこれらの間には、「密接な因果的結合が存在するからである」(ibid.) と。

しかし、だとすると、あなたが生まれてからずっと水槽脳として生きてきたというのではなく、あるとき突然に脳手術を受けて水槽脳にさせられたという状況、つまり冒頭でパトナムが示唆した状況の下で、しかも記憶が無傷のままに存続している場合、水槽脳であ

るあなたは、水槽脳になる前の本物の樹木や手術室や水槽を指示できるということにならないだろうか？ つまり、このときのあなたの言明「私は水槽脳である」は、自己論駁的どころか、まさしく真なる言明であるということに。というのも、そのときまでのあなたと本物の樹木や手術室や水槽との間には、確かに確固とした因果的つながりがあったからである。例えば、神山和好は「パトナムの証明が成立するためには、われわれは生まれながらに水槽の中の脳である、という設定が必要である」と述べる（神山 [2004] 32）。また、デローズも、生まれつきの水槽脳は水槽-日本語において一般に真なることを述べるというパトナムの主張に対して、ごく最近に手術された水槽脳はほぼ全面的に偽なることを述べるように欺かれる、と示唆している（DeRose [1995] notes 2）。

さて、本稿が論ずるパトナムの議論の一局面とは、まさしくこの問題、つまり十分長い間ふつうの経験をした後で、ごく最近、水槽脳にされた人物は、過去の記憶によってこれまでの事物を指示することができるのだろうか、という問題である。本稿の結論は、いかにその過去の記憶が鮮明であろうとも、また手術がたかだか数分前のことであろうとも、水槽脳としての経験を開始した瞬間から、それ以前の事物を指示することはできない、というものだ。

## 2. 記憶喪失と世界喪失

水槽脳はその記憶によって手術以前の世界の事物を指示することはできない（語ることはできても）、という結論を引き出すためのモデルは、指示に関する基礎的な二項関係に求めることができる。それは、きわめて当然のことながら、指示する項目と指示される項目の二項の存在であり、この場合は前者が記憶（表象）、後者が世界内の事物である。この二項関係は、どちらか一方が失われれば崩壊し、指示関係も成り立たなくなる。記憶喪失の場合は、後者の世界内の事物が依然として存在しているのに対して、前者の指示する項目の側が失われるケースだ。この場合、当人から指示する能力が失われたことによって、例えば当人が所持していた事物への指示は不可能になる。では、水槽脳の場合はどうか？ この場合には、仮定により手術後も当人に記憶は保持されており、指示する能力は失われていない。しかし、指示される側、すなわち世界が丸ごと失われたために、二項関係は崩壊し、以前の世界内の事物への指示は不可能となる。つまり、水槽脳の場合は、記憶喪失とパラレルな世界喪失が生じていると考えるべきである。

この結論は、指示における因果的結びつきに、世界内在的な視点の徹底化からある種の制約を課すことによって導かれる。ここで世界内在的な視点とは、「経験の主体はある特定の世界にのみ存在し、複数の世界に同時にまたがって存在することはできない」というものだ。これは、水槽脳となって経験を開始している人物  $a$  は、彼の経験する世界  $w_1$  に属し、 $a$  の経験世界を生み出している世界（ $a$  の元の脳を操作している世界） $w_2$  に同時には存在しない、という控えめな主張をしているにすぎない。ただし、ここから、 $a$  にとって自分の脳は（自分の身体の一部として） $w_1$  に存在するが、脳操作を行っている者たちにとつ

ては  $a$  も  $a$  の脳も  $w_2$  に存在する、ということが帰結する。

さて、世界内在的な視点を指示関係に関して一貫させるなら、次のような因果説的な制約を指示に関して要請することは理に適っているだろう。それは、ある世界内における指示する項目は、心的表象であれ言語記号であれ、その世界内でその項目と因果的つながりをもつものしか指示しえない、ということである。もちろん、他の可能世界に存在するものについてはいくらかでも語ることはできるし、語るだけならば、われわれは矛盾や不可能なこと、さらには不可能世界についてすらも語ることはできる。ただし、ここで提案している〈世界内在的な指示〉からすれば、われわれは他の可能世界に存在するものを指示することはできない。ついでに言えば、この指示概念からすると、クリプキの固定指示詞 (rigid designator) は、あらゆる可能世界で同一の個体を「指示」するのではなく、同一の個体に関して様々な様相的な「語り」をするための枠組みの一部である。

したがって、先のパトナムの水槽脳による思考「眼前に樹木がある」における「樹木」の指示を、「イメージ樹木」や「電子工学的インパルス」や「プログラムの特質」と解釈するのは、あくまでその水槽脳となって経験を開始している当人ではなく、その脳を操作している者たちである。というのも、眼前に樹木を見ている当人  $a$  による「樹木」の指示は、彼が存在する  $w_1$  の目の前の樹木に届くのであって、 $w_2$  のいかなる事物にも届かないからである。一体  $w_1$  のどのような因果連鎖を辿ったら、 $a$  は  $w_2$  の世界に到達できるというのだろうか？ この点は、大森が描いた脳操作の第二のシナリオ、記憶を保持したままでの脳操作において鮮やかに示されている (大森 [1981] 184-188 頁)。そのストーリーでは、手術後に異国らしいどこかの国のどこかのホテルで目を覚ましたあなたは、生々しい記憶を頼りに、「操作されているはずの自分の脳」に辿り着こうとして故国日本のあのルネ・デカルト病院の手術台を目指す。しかし、うまくその病院に辿り着き、偶然にもあの同じ手術台とおぼしきものを発見したとしても、それらは、元のルネ・デカルト病院でも元の手術台でもない。それどころか、その手術台の上に自分自身をあなたが発見したとしても、それは本当のあなたではない。あなたが二人存在することはないからだ。あなたが発見したのは、元の病院やあなた自身と質的に同一の何かであっても、数的に同一のそれらではない。脳操作されているあなたは、現在のあなたにとっては端的に消失したのであって、それが存在しているのは、脳操作している世界の中であり、その世界に存在する者たちにとってである。したがって、手術後に経験を開始する  $a$  が指示しうるものは、少なくとも  $a$  の存在する世界  $w_1$  に存在していなければならない。

しかし、過去世界はある意味でもはや存在しないとも解釈されるのだから、以上の議論は過去についてのわれわれの指示を同じように脅かしはしないだろうか？あるいはそれは、夢の内容に対するわれわれの指示や、世界  $w_2$  において水槽脳の経験内容が事細かに操作されているという想定と整合的なのであろうか。

### 3. 過去世界への指示と夢世界への指示

まず、過去の事物や出来事への指示について考えてみよう。あなたが普通に朝目覚めるとき、あなたは昨日のことやさらにそれ以前のことを記憶している。そしてその記憶に従って、昨日の夕方に食べた居酒屋の刺身や数ヶ月前の母校主催のフォーラムのことについて友人に話す。その際、あなたが過去のこうした事物や出来事を正しく指示しているのは間違いない、としよう。他方、あなたは手術後に水槽脳として初めて目覚める。あなたは手術前の、昨日の夕方の刺身や数ヶ月前のフォーラムのことを思い出し、暢気にもそれらについて友人に話し始める。この二つの状況は、ここまではまったく同じであるように見える。そしてどちらの場合も、目覚めたときにはすでに過去世界はある意味でもはや存在しないのであるから、この両者にどのような違いがあるというのだろうか。

しかし、両者の違いはある意味で歴然としている。というのも、水槽脳の場合、先ほど触れた論者らによれば、あなたの記憶表象に求められているのは、あなたの今の脳とともに世界 w1 に存在する昨日の刺身や数ヶ月前のフォーラムへの指示ではなく、あなたの元の脳とともに世界 w2 に存在するそれらへの指示だからである。しかし、これはわれわれの<世界内在的な指示>によれば不可能である。普通の場合にあなたの記憶表象の因果的連鎖を辿れば、それらを因果的に生み出す原因となった昨日の居酒屋や数ヶ月前のフォーラムの会場に行き着くだろう。つまり、それらからの因果的連鎖のつながりが、あなたの現在の脳の記憶表象へと達しているだろう。では、水槽脳の場合はどうか。この場合も、あなたには普通の身体と脳があり、あなたの脳に存在する記憶表象の原因を辿れば、世界 w1 に存在する昨日の居酒屋や数ヶ月前のフォーラムの会場に行き着くはずである。つまり、水槽脳の操作は世界を丸ごと生み出し、その世界の中には過去世界も含まれているのであるから、過去からの名前の因果的連鎖を引き継ぐことによって織田信長やジュリアス・シーザーのような歴史上の人物を現在われわれが指示するように、水槽脳が経験する世界 w1 においても、過去への指示は<世界内在的に>行われているのである。したがって、手術に関する水槽脳の記憶すらも、手術後に出現する世界 w1 で値踏みされねばならない。それゆえ、あなたがこのときなおも世界 w1 に存在する事物に満足せずに、「自分の本当の脳はこれではない、脳手術のことを私は間違いなく記憶しているのだから」と言うとするなら、それに対する答えはこうである。「その話が本当であるなら、あなたの元の世界 (w2) はいまや消失したのであり、したがって、あなたの記憶はその世界に存在するなにものも指示することはできない」。

仮に脳読み取り装置の技術が向上し、脳が蓄えている記憶を取り出してスクリーンに映し出せるようになったとしても、映し出されたその光景はすべて、水槽脳であると主張するあなたの世界 w1 に存在する機材によってその世界内に作り出された、その世界内の映像にすぎない。それは当然である。その記憶をもった脳は、世界 w1 に存在する脳だからである。

では、あなたに水槽脳の手術を施し、それを操作している世界 w2 の住人は、水槽脳であ

るあなたの経験内容、あるいはあなたの経験する世界 w1 に存在するものを指示できるのだろうか？ パトナムのストーリーでは水槽脳の経験をコントロールするのは超科学的コンピュータだということになっていた。しかし、それらを駆使する科学者の集団や、それを見守る観察者の集団を想定するのは何らおかしいことではないだろう。彼らは世界 w2 に存在しながら、あなたの世界 w1 に存在するものや出来事を指示できるのだろうか？ 実は、この想定では世界 w1 はある意味で世界 w2 に寄生する形で存在しているのであるから、世界 w2 は世界 w1 を寄生的に含むと言っていいだろう。そして、この種の寄生的な内含は、寝ている人の夢の世界がその人（夢主）の属する世界に含まれるのと同様に、それほど奇怪なことではない。

そこで、夢の場合を考えてみよう。夢の内容が夢を見ている当人(夢主)への刺激によってかなり影響を受ける、というのはよく知られた話である。例えば、救急車のサイレンが鳴った夜の当人の夢には、サイレンを鳴らす救急車が何らかの形で登場したりする。つまり、夢主の存在する世界から夢の世界への因果的なつながりは、現在のところ、コントロールというほど意図通りのものではないにせよ、一般的な因果的影響を与えうるとい程度には確かなものだ。ここで、先ほどの脳読み取り装置によって、夢主の脳内の神経活動から夢主が経験しているであろう場面を、枕元のスクリーンに映写することができるなら(原理的にはできるだろうが)、夢主の傍らにいるわれわれはその映像の中の事物や出来事を、ちょうど、テレビ映像に映る外国の街角のシーンを「指示する」というのとほぼ同じ意味で「指示する」ことができるだろう。その際、夢主の脳活動を指示することが、すなわち、夢の世界内の事物を指示することに他ならない。件のスクリーン上の映像への指示は、脳活動への指示の映像的なパラフレーズにすぎないであろう。ただし、極めて重要なポイントだが、その逆は成り立たない。夢の世界の登場人物が夢主の世界に存在する事物や出来事を指示することは、<世界内在的指示>の概念からして不可能である。というのも、夢の世界のなかのいかなる事物、いかなる「指示する項目」に関して(たとえ、それが夢の中の登場人物である「私」の脳状態であろうと)、どれほどその因果連鎖を辿っても、夢主の現実世界に到達することは不可能だからだ。唯一ありそうな話として、その連鎖の行き着く先を夢の世界の因果的な原因である夢主の脳としてみたところで、夢の世界の時空的全体を創りだしている夢主の脳が、夢の世界の時空のどこかに「へその緒」のような仕方で存在しているわけではない。この意味で、<世界内在的指示>の関係は、世界間で一方通行的である。もちろん、現実世界から夢の世界への一方通行的な指示が可能だとしても、どちらの世界に存在する事物も他方の世界に物理的に持ち込むことはできない。現実から夢の世界への因果的なつながりは、夢主の脳という一種の<特異点>においてのみ可能なのである。つまり、世界内在的な指示関係を可能にする因果的なつながりは、夢や水槽脳の場合、脳という<特異点>への因果的な関与という仕方でのみ可能であり、しかもそれは一方通行的にしか可能ではない。

#### 4. 水槽脳の世界への指示

さて、あなたに水槽脳の手術を施した人々やその見物人たちは、あなたが経験する世界  $w_1$  に存在する事物や出来事を指示することができるのだろうか。これは、超科学的コンピュータがあなたの経験のすべてをコントロールする、というパトナムの想定からして可能だと考えられている。そして事実、状況は、前節で見た夢の場合と基本的に同じだと考えられる。ただ、この場合、夢の場合のような漠然とした因果的影響どころか、正確無比の因果的制御が行われているという違いがある。水槽脳が経験する世界  $w_1$  は、先にも述べたようにそれを操作する側の世界  $w_2$  に寄生的に含まれる。したがって、操作される水槽脳という<特異点>への指示が、その寄生的世界  $w_1$  への指示となる。未来の脳読み取り装置を使うなら、世界  $w_2$  の人々の前で、水槽脳であるあなたが経験する光景が経験主体であるあなたの視点から映像化されるだろう。もちろん、その映像の素材はすべて世界  $w_2$  のものであり、世界  $w_2$  に存在する人々は、その映像に映る事物や出来事を、あくまで世界  $w_2$  に存在するスクリーンや電子的映像に対する指示として行うことができる。しかし、世界  $w_1$  に存在する事物や出来事への指示として考えるなら、むしろ夢の場合と同様に、その指示は世界  $w_2$  に存在して操作されているあなたの脳への指示によって果たされるというべきであろう。

ちなみに、この<世界内在的指示>の概念からすれば、いわゆる可能世界に存在すると想定される事物や出来事は、厳密な意味では「指示」できない。われわれは、それらについて「語り」うるのみである。もちろん、それらの語りが多岐にわたる哲学的洞察をわれわれに数多くもたらすとしても。また、この指示概念は、確定記述に関しては、それを指示句とするのではなく存在命題の一部とするラッセルの解釈に基本的に従う。「2088年の日本の総理大臣」という確定記述句によって、われわれは何ものも指示してはいない。

最後に結論に代えて、再度、水槽脳が経験する世界  $w_1$  (および夢の世界) と水槽脳を操作している世界  $w_2$  の関係をまとめておこう。

(1)  $w_1$  の住人は  $w_2$  のいかなる事物も出来事も指示しえない。 $w_1$  の住人  $a$  が脳操作直前までの  $w_2$  の記憶を持っているとしても、それに対応する  $w_2$  の存在物を指示しえない。なぜなら、 $w_1$  の住人にとって世界  $w_2$  は端的に消失したからである。もちろん、 $w_1$  の事物を  $w_2$  に物理的にもちこむことはできないし、その逆でもある。

(2)  $w_2$  の住人は  $w_2$  に寄生して存在する  $w_1$  の事物や出来事を指示しうる。しかし、その指示は、操作される水槽脳という<特異点>への指示に他ならない。水槽脳が経験する内容の  $w_2$  における解釈(脳読み取り装置による画像化や音声化)はすべて  $w_2$  の事物へのパラフレーズであり、水槽脳への指示のガイドとして役立つにすぎない。また、 $w_2$  の住人が夢を見る場合、その夢の世界もまた  $w_2$  に寄生的に存在する世界として、 $w_2$  の住人が指示することができる。もちろん、夢主以外の他人が、夢主の夢の世界の住人としての経験を持つことは不可能である。

(3)  $w_1$  の住人が夢を見ることもあろうし、 $w_1$  の住人  $b$  に対して水槽脳の手術を施すことも可能であろう。その場合、水槽脳である  $b$  の経験する世界  $w_0$  と  $w_1$  の関係は、 $w_1$  と  $w_2$  の関係と同じである。世界  $w_0$  の住人は世界  $w_1$  を指示できないし、ましてや  $w_1$  を因果的に生み出している(元の水槽脳の)世界  $w_2$  を指示することもできない。

(4) 水槽脳  $c$  を操作する世界  $w_2$  が実は世界  $w_3$  に存在する水槽脳  $d$  の操作によって出現していた、ということも可能であろう。その場合も、 $w_2$  と  $w_3$  の関係は、 $w_1$  と  $w_2$  の関係と同じである。そうではなく、もし仮に世界  $w_1$  の住人  $a$  が、その世界から出発して、世界  $w_2$  とその世界に存在する水槽脳  $b$  を指示することが可能であり、「私は水槽脳である」という言明が真となることがあるならば、 $n$  個の水槽脳が次々に操作される多重水槽脳仮説の元で、 $a$  の発言した「水槽脳」はどの世界に存在する水槽脳を指示するのかをどうやって決定できるのだろうか？ 世界  $w_1$  に存在する形而上学的実在論者と、少なくともく世界外在的指示概念の信奉者は、この問いに答えなければならない。われわれの現実世界の現に経験されている状況が世界  $w_2 \sim w_n$  ではなく、世界  $w_1$  と同じである（つまり、この世界には実際に操作されている水槽脳は存在しない）ことを考えると、先の形而上学的実在論者も件の信奉者も、それに答えるすべはないように思われる。

注：

(1) 本稿は、金沢大学哲人文学類学研究室の平成 23 年度卒業生、片岡雅知君からの講義における質問と、それに続く彼との数度に渡る議論から生まれたものである。

参考文献：

- 大森荘蔵、1981、「夢みる脳、夢みられる脳」『流れとよどみ』所収、産業図書
- 神山和好、2004、「水槽の中の脳型懐疑論を論駁する」『科学基礎論研究』Vol.32, No.1
- 津留竜馬、2000、「パトナムのモデル理論的議論と水槽の中の脳」『哲学誌』42号
- Brueckner, A. J., 1986, *The Journal of Philosophy*, Vol.83, No.3.
- DeRose, K., 1995, "Solving the Skeptical Problem," in *Epistemology: An Anthology*, 2nd ed., 2008, E. Sosa, et al. (eds.), Basil Blackwell.
- McIntyre, J., 1984, "Putnam's Brains," *Analysis*, Vol.44, No.2..
- Putnam, H., 1981, "Brains in a vat," in *Reason, Truth and History*, Cambridge University Press. パトナム、H. 「水槽の中の脳」『理性・真理・歴史』野本・中川・三上・金子訳、法政大学出版局、1994.
- Tymoczko, T., 1989, "In Defense of Putnam's Brains," *Philosophical Studies*, 57.