

へわたしは脳に操られているのか
意識がアルゴリズムで解けないわけ

【立ち読み】

はじめに 人間に自由な意志はあるのか 8

第1章

人を殺したの
は脳のせい？

13

意識にのぼらない計算／道徳的責任を問う根拠

第2章

意志はころがり落ちる石なのか

23

神経科学者は悪魔か／決定論とランダム性

第3章

二つの対立する答え

36

女神ヘラの呪い／ほかに取りうる選択／対決すべき問題

第4章

頭のなかの嵐

50

ジャン・ヴァルジャンの自問／限りのない問題／

意識のない神経系から、いかに意思決定力のある主体が生まれるか

第5章

抑えられない衝動

64

いきなり友人を罵倒する／やめられない人々／「私」障害

第6章

神経科学者の見解は間違っている

77

「意志の座」は前頭葉なのか／健康な脳が健康な行為を決定するとは限らない

第7章

理性は情動に依存する

84

なぜ最悪の選択をしてしまうのか／それでも最終的な発言権はある

第8章

決断の引き金が明らかに 98

壊れたブレーキと事故／準備電位と自発的行動／「拒否する」能力は残る／
リベットの試験への批判

第9章

マジシャンとしての脳 113

錯覚をつくる三つの要因／意志と行動を生むシステムは別なのか／
行動をコントロールしているという錯覚／ウエグナーの説の問題とは？

第10章

心や体の動きを予測する 132

行動を予測するテクノロジー／心を読み取る機械／こっそり下見しているだけ

第11章

人間はプログラムされたマシンか 150

心を変える薬／脳の革命

第12章

悪徳の種が脳に植えられている？ 163

衝動的暴力・計画的暴力／ホルモン刻印／囚人になる運命

第13章

倫理の終わり 182

二つの心、二つの自己／信念を形成する装置／ガザニガらの主張は正しいか

第14章

意識の深さを探る 195

閉じ込め症候群／自由意志は非科学的？／決定論のどこが誤りなのか

第15章

アルゴリズムは「限りのない問題」を解けない

213

テロ対策のジレンマ／経験と熟考／私たちは決定論的システムではない／人間の思考とアルゴリズムのちがいは？

第16章

内面世界を意識的に旅する

233

最も妥当な筋書きを求めて／思索的内省／アルゴリズムから内省への移行／脳はいかに無限の意識を生み出すか

第17章

道徳的行為主体はいかに生まれるか

259

決定されていないが、ランダムでもない／脳とカオスと量子力学

第18章

心の宮殿

269

心の深みをうまく進む能力／脳が殺人をさせたわけではない／倫理の基礎／進化の勝利

謝辞
284

もっと詳しく知るために (1) 注 (9) 参考文献 (24) 解説
324

はじめに 人間に自由な意志はあるのか

フォードル・ドストエフスキーは、一八六四年の著書『地下室の手記』（光文社など）で、名前は一度も出てこないが、自意識に悩まされている男のことを語っている。男は、自分は優柔不断で何もできなくなる意識の病にかかっていると考えている。意識過剰で、自分の考えや意図を気にしすぎ、自己批判的すぎるといふのだ。さらに悪いことに、自分の決断は社会の期待によって操られ、自分の体の機械的な力にも支配されているという被害妄想を持っている。この本は回想録として書かれているが、「地下室の住人」が自分の思考と行動をコントロールしていると納得できるように行動しようとしている。奮闘しながら、自分の意志を世界に行使しようとするときの、底知れぬフラストレーションを表現している。

彼は自分のほうに向かって歩いてくる人に、道を譲りたくないことを表明するためにわざとぶつかる。人の予想を裏切るために、歯痛は快楽だと主張する。内心では親しくなりたいと思っている昔の学友の夕食会に参加し、おおっぴらに彼らを侮辱しようと思心する。ほかの客は彼がそこにいることを望んでいないとわかっているが、復讐についてあれこれ考えながら優柔不断に歩きまわり、帰ろうとしない。自分がどうふるまうかを他人が予想していると感じるたびに、彼は意地悪するためだけに予想とはちがう行動をとる。

ある日、地下室の住人はリーザという娼婦と出会い、彼女との会話に引き込まれる。そしてとくに理由もなく、彼女の職業が受ける屈辱について説教することにする。地下室の住人はリーザが自分の言うことに心を動かされているのを知って驚き、彼女に影響をおよぼすことができるという期待に胸躍らせる。リーザは彼のひどい皮肉に辛抱強く応じ、地下室の住人は次第に彼女を好きになっていく。しかしその気持ちが彼を怒らせる。その気持ちのせいで二人の交流がもどかしいほど複雑になり、地下室の住人はその気持ちが自分ふるまいに影響することを警戒するのだ。そのため、彼はひとりつきりなしに彼女を馬鹿にして、とうとう彼女は考えにふける彼をひとり残して立ち去る。

それでもリーザは彼のふるまいに寛容で、彼のもとを訪ねてくる。地下室の住人は、彼女は自分のことを愛情深く見てくれる唯一の人、いつの日か自分を理解してくれそうな唯一の人かもしれないと気づく。彼がこの新たな人間関係に力を注ぐべきなのは明らかだ。しかし、これが彼のなすべき決断——彼の立場にある分別のある人なら誰でもする決断——だったからこそ、地下室の住人は意地悪くなくて、ちがう行動をとることにする。彼は軽蔑のしるしにリーザにお金を投げ、屈辱を与え、彼女は永遠に去ってしまう。地下室の住人は内なる心の葛藤にもどり、再びシニカルな取りとめのない話を始める。

本書の題辞（12ページ）は地下室の住人の手記から引用している。¹ドストエフスキーの小説のなか

地下室の住人は終始、自分が自分の行動をほんとうにコントロールしているかどうかという問題に取り組み、自分は純粹に機械的な世界（オルガン）の取るに足らない部品（「オルガンの音栓」）にすぎない可能性をつねに突きつけられている。さらに、自分について十分な情報を持つている誰かが、自分のあらゆる決断を予測するために公式のようなものをつくる可能性があることに不安を覚える。そのため、彼は必然性の手中から逃れたいと思つて、とつびな行動をしようとする。自分で選んでできる能力、周囲に影響を与える力が、自分にはあるのだと示すことだけを追い求める。

地下室の住人が提起している問題は、一八六四年より現代のほうが現実的な意味を帯びている。実際、それは私がずっと悩んできた問題だ。神経科学分野の発展のおかげで、人間の行動が脳内のニューロンとそれに関係する化学物質との相互作用にどう関係しているか、理解がどんどん深まっている。現代の科学者の大半は、このような生化学的実体の仕組みが、人間の行動にとって唯一の決定因子であるという前提で研究している。周知のように、生化学的相互作用は一連の厳格な法則、つまり化学と物理の法則にしたがつているので、数学的方程式のかたちで表現できるほどである。私たちの思考や選択が脳内の化学物質によつて引き起こされるのなら、やはり物理の法則で決定しているはずであり、したがつて原理的には数学的方程式で表現できる可能性がある。

これこそまさに、地下室の住人を恐怖に陥れている可能性だ。選択の自由がなければ、彼は人に（結果の良し悪しにかかわらず）行動の責任を負わせることはできない。なぜなら、人は自由・自由・選べないからだ。選択の自由がなければ、地下室の住人はルールにコントロールされている自然界の駒

にすぎないので、世界に対する影響力があるとは主張できない。それゆえ彼は悪意に満ちた人間になり、誰も予測できないと思うような行動をとるのだ。

生物学的メカニズムが私たちのすることすべてをコントロールしているというのが真実なら、私たちはどうなるのだろうか？ それはつまり、私たちは機械であり、オルガンの音栓であることを意味する。

私としては、けつしてオルガンの音栓にはなりたくない。この気持ちから、この苦境を抜け出す道を探そうと奮起している。自由意志と道徳的責任に対する神経科学の脅威は大きいが、和解の道はある。脳科学の手法や発見に対する信頼と、意図を持つ道徳的行為主体としての人間の尊厳に対する信念、両方を維持できる方法がある。少なくとも、私はここでそれを論証しようと思う。本書でこれから、そのような和解がありえることを主張し、これらの見解がどうつながりえるかについて考えを示すつもりだ。

本書は大きく三つに分けられる。第一部（第1章〜第4章）は、ここでざつと紹介した問題の意味合いの深さを探る。第二部（第5章〜第13章）では、私たちの自由意志と道徳的責任を脅かすように思われる神経科学の鮮やかな発展の概要を説明する。そして第三部（第14章〜第18章）では、考えられる解決を示す。全体として本書を、たんに私が長い論理の道筋を提示する舞台にするのではなく、自由、倫理、そして人間の意義にまつわる根本的問題を探究する書にするつもりだ。

人工知能のめざましい進展は、人間の能力に迫り、さらに超えつつあるようにさえ見える。やがて、心や意識をもったAIも生まれると予測する研究者も少なくない。私たち人間と区別のつかないようなアンドロイドが、未来に登場するのだろうか？ だが、想像の翼を広げる前に、あたりを見回してほしい。いまあなたの周りにいるひとたち、そしてあなた自身ですら実は「アンドロイド」だったとしたら？

これはSFの話ではない。脳科学の成果は、実際、私たち自身が〈脳〉に操られるアンドロイド（生きたマシン）にほかならないというのだ。たとえば、本書でも紹介されているベンジャミン・リベットの実験。この実験は、私たちが意識的に自覚する前に、すでに脳が行動しようと実行しはじめていることを明らかにした。つまり脳（無意識）が私たちの行動を、まず先に自動生成しているというわけだ。また、脳画像をモニターすることによって、意思決定や体の動きを事前にある程度予測できるといふ別の実験結果からも、このことが伺える。

ほかにも本書には、脳が私たちの心や行動を操っている研究や事例が次々と登場する。脳障害による人格の変貌、心を変える薬、衝動を抑えられなかつたり、暴力や犯罪をおかしやすい脳の傾向、妄想を確かな事実として信じてしまう症状など……。想を確かな事実として信じてしまう症状など……。想を確かな事実として信じてしまう症状など……。

今日の脳科学では、脳が人間の思考や行動を因果的に決めており、自由な意志は存在しないと「決定論」が多勢を占める。そしてこのことは、私たちの社会や倫理にとつても見過ごせない大きな問題を含んでいる。たとえば悪事を働いても「それは脳のせいだ」ということになってしまつてからだ。現に犯罪者を裁く法廷でこうした主張がなされることもあり、極刑を免れた例もある。

決定論と対立するのが、「自由意志」はあるという立場であり、本書もこちらに属する。すなわち、脳による無意識の決定を超えて、人間には意識的に熟考する能力があるとする。一方、決定論でも自由意志はあるという「両立論」も主張されるが、本書はこの説では疑問に答えられないと退ける。また、心と脳（体）は別だとする二元論の立場もとらない。

では、「自由意志はある」という根拠はなにか？ 本書は意志と脳にまつわるさまざまな研究や事例を参照しつつ、その根拠を批判的に探っていく。脳科学・認知科学における自由意志の主な論点を手際よくまとめられており、読者はいかにこの問題が広範な領域に及ぶかに気づかされるだろう（なお、心理学・哲学的な「自由意志」の論点については、『社会心理学研究（Vol. 31・No. 1）』所収の「自由意志信念に関する実証研究のこれまでとこれから」に詳しい。この論考はネットから無料でダウンロードできる）。

また決定論者でも、脳神経倫理学を主導するガザニガのように、最終的には自由意志は残るとの主

張もある。しかし、こうした見解も本書は批判しており、脳科学の主流派にいかにか斬り込んでいくかも読みどころとなっている。

とはいえ、脳じたいが決定論的システム（アルゴリズム）を基盤とすることは認めざるを得ない。それを認めたくえで、なおかつ自由意志が生まれることを明かさなければならぬ。まず本書は、決定論のロジックには「飛躍」があり、理論というより「世界観」に過ぎないことを説明する。さらに、「限りのない問題」というキーワードによって、人間の論理的思考はアルゴリズムではないことを明らかにする。「限りのない問題」とは、たんに定義付けできないだけでなく、関係する概念がわかってても、その解釈の方法も無数にあるような問題であり、アルゴリズムでは解けない。しかし人間は、「アルゴリズムを超越する——状況を理解し、意味を認識し、想像し、意識的に熟考し、限りのない問題を論理的に考え、自由な行為主体として行動する」ことができるのだ。

では、どのようにこうした自由な行為主体が、脳から生まれるのか？ 本書は興味深い説を展開していくが、そのひとつは脳科学者アントニオ・ダマシオの「ソマティック・マーカー仮説」に想を得ている。ソマティック・マーカーは、知覚的な刺激と、過去の情動や記憶・意識的な経験のあいだをつなげる。そして、その連携は、脳内の前頭葉が担っている。これに似て脳は、まず知覚や経験のデータ処理をアルゴリズムとして行いながら、内省的な思索へと移行していくのではないか。

さらに、自由意志や意識をもつ行為主体は、決定しているプロセスとランダムなプロセスとの相互作用によって、脳内で生まれるのかも知れない。本書はカオス理論なども援用するが、その解明には新たな科学的アプローチ（意識の物理学）が欠かせない。

意識（自己認識）をそなえた人工知能を創造する難しさも、まさにこの点にある。従来のアルゴリズムの能力を高めていく方向だけでは難しいのだ。しかし、人工意識を目指す研究は日々進んでいる。たとえば、人間の認知発達や生命進化を模した学習プロセスを組み込む方法がある。赤ん坊が認知発達で得るような客観的な視点を獲得し、自発的な意志によって、自ら考え、学習し、限りなく成長しうるシステムが構想されている（荒木建治ほか『心を交わす人工知能』などを参照）。また、脳型コンピュータと呼ばれるニューロモーフイック（神経形態学的）なプログラムが、やがて生命のように「カオスの縁」で働き、ある種の自己認識を持つようになることも考えられる（ジョージ・ザルカダキス『AIは「心」を持てるのか』などを参照）。こうした未来像も含めて、本書には「自由と倫理」「意識とアルゴリズム」をめぐる刺激的な洞察があふれている。

最後になったが、脳科学・心理学・コンピューターサイエンス・哲学・文学など、広範な領域を横断する原著を、わかりやすい確かな訳文にいただいた大田直子さんに多大の感謝を！

本書出版プロデューサー 真柴隆弘