

2012 年度夏学期 金曜 4 限

哲学 I (石原孝二教官)

試験対策まとめ

[はじめに]

全 12 回の講義を、スライド資料・講義中のメモ・ネット情報の拾い読みをふんだんに使って焼き上げました。開封後はお早めにお召し上がりください。

[試験について]

- 知識を問う設問：12 問程度(計 60 点程度)
「～～とはどのようなものか。～行で述べよ。」のようなもの。短め。
- 議論を求める設問：2 問程度(計 40 点程度)
意識と心全般に関する論述問題。一定の知識を前提とする。

[試験対策について]

- 講義資料を読み返す
- あきらめる
- どうしてもという時はこのまとめを読む

ハッキリ言って講義内容は難しいので、必要でない方は素直に撤退したほうがいいと思います。このまとめもそれほど分かりやすくないです(太字すら使っていない)。誤字・誤認識・誤信念についてはお許しください。

【0. 意識の謎】

[概略]

本講義は、「意識」及び「心」とは何かを問う歴史を概観するものである。本章では、その前段階として、「意識」及び「心」とは何か、だれが持ちうるのかについて簡単な考察がなされている(読まなくてもいいと思う)。

「心」とは何かに関する問題は、古代より世界各地に存在した(インド哲学、ギリシャ哲学など)。その中心となる論点は、「心」と「身体」との関係の問題であり、これはデカルトなど近代哲学まで続く論点でもある。

他方、「意識」の概念が登場したのは、デカルト以降の近代哲学においてである(ということにしておく)。この論点はフロイトなどを経た後に脳や神経系に着目した自然科学との交わりを深め、現在では神経科学などとの関わりが深い。本講義の中心はこの、哲学と自然科学との結びつきが顕著に見られた20世紀中盤～後半以降の「意識」及び「心」についてである。

さて、「意識」は近代初期においては人間特有のもののみなされる傾向も存在したが、現在においてはどうか。他の動物、あるいはロボットや人工知能などに「意識」は存在するのか。その研究例として講義内では、ブレイスウェイトの『魚は痛みを感じるか?』(2010)が紹介されている。ブレイスウェイトによれば、意識を構成する要素は3つ存在し、実験の結果、魚に関してはその3つをすべて持っていることが実証された(らしい)。しかしここで重要なのは、(実験対象である)魚が「意識」を持つかどうかという事よりも、そこで定義されている「意識」とは何か、である。

[補足説明]

●意識の3カテゴリー

・現象的意識 P-consciousness

自分のまわりの出来事を直接的に感じ取ること。ブレイスウェイトは、魚が情動処理に特化した脳領域を備えていること、また周囲の状況に応じて回避すべき対象に対する見方を変化させていることなどから、魚が感覚力、すなわち「現象的意識」を持つことを示している。

・アクセス意識 Access-consciousness

現在、あるいは記憶に基づく心の状態について考え、または描写すること。この能力を駆使して心的表象を形成し、主体は行動(=反応)を決定する。換言すると、反応に基づいた意識であり、より高次のレベルで利用可能であるもの。魚については諸々の問題解決の際に、記憶に基づいた対応など、心的表象を活用する能力が見られたことから、ブレイスウェイトは魚が「アクセス意識」を持つと認定している。

・モニタリングと自己意識

3つの中で最も高次の意識活動と思われる。自分自身の行動を心の中で実行し、それによって生じる状況について思い浮かべることで、様々なシナリオを考慮する能力。ブレイズウェイトは同著書内で、ハタとウツボの関係を例に挙げてこの能力について説明している。ハタは獲物を前にして、単独では仕留められないことを予測して、一旦その場から泳ぎ去ることでウツボの協力を引き出そうとする。するとハタの見込み通り、ウツボは獲物のもとへ近づいていく。ここにおいて、ハタは、自らの行動を心の中で省みることによって目的達成への道筋を立てていることが伺える。

【1. 表象と意識】

[概略]

前講で述べたように、「意識」に関する議論はフランスの哲学者デカルトから始まる。デカルトは方法的懐疑を採用したうえで、あらゆる疑わしいものを排除したのちに最後まで残るものとして、「観念そのもの」への自己の考察を発見した。自己の中に浮かび上がる「腹が減った」という表象の内容は、誤りであることが十分ありうる。しかし、「腹が減った」という表象そのものを疑うことはできない(何故ならば実際にそれを自己が感じ取っているから)。この事から発展させて、デカルトは「精神」と「身体」の二元論における精神の優越性、不滅性を主張した。

イギリスの哲学者ロックは、これに対抗する形で、意識を「人格」と結びつけて考察した。人格は、主体が「時間と場所を異にして同じ」であるという自己同一性の基礎となるものであり、しかしそれは記憶(=意識)によって支えられているものである。記憶の及ばない領域については人格同一性が認定されず、その意味で「意識」は精神に基づく不滅なものであるというデカルトの思想とは異なり、主体の経験を重視している。

このイギリス経験論を発展させた哲学者がヒュームである。ヒュームはロックの議論を逆転させて、自己の同一性の存在も、「観念そのもの」の存在も否定している。このような主体に「同一性」の幻想を持たせているものが、個別の表象であり、記憶(=意識)であるという形で、意識について考察を行っている。

デカルト的な大陸合理論と、イギリス経験論を統合して構築されたものが、カントに由来するドイツ観念論である。カントは、全ての思惟に先立って与えられている表象を「直観」として、個別の表象である経験的統覚と区別して、これらを統括する直観を「純粹統覚」と名付けた。

講義内における近代哲学者の取り扱われ方が若干まとめづらい。以下では、哲学史的な定義ではなく講義内でどのように扱われたかに特化して、補足説明を行う。

[補足説明]

●デカルト『省察』、『方法序説』

デカルトは、自己をモノに向けて対象化して、そのモノのうちに自己の観念を見て取る行為、つまりモノを経由して表象を照射する行為を「想像」と称して、これについては「間違えることがある」として排除した。一方、自己を自己自身に向け、精神そのものに内在する観念を考察する行為は「理解」であるとして、これは「間違えることがない」とした。

前者は、指をナイフで切ったから痛い、体が火照って暑いなど、自らの身体に基づくものも含めたすべての「モノ経由」の表象である。後者は精神のうちに完結した行為であり、ここで後者を疑えないものであると判断したことからデカルトは、「身体」に対して「精神」が全く独立した(優越する)本性を持つ別の存在であると考え、死滅する「身体」に対応させて「精神は不滅である」と主張した。

●ロック『人間知性論』

社会契約説についても別な著書で論じているロックは、人間の社会の中における立ち位置を決定するものが「人格」であるとした。社会の中で賞罰を受ける対象となっているものも「人格」であり、それ故にたとえば人格の同一性が保証されない「心神喪失」などの場合においては、罪が問われないなどの制度も存在する。

ロックにおいては、記憶に残っていない意識経験は人格の同一性を構成しないため、そのようなまとまった意識経験はまた別個の人格として認識される(これは解離の症状観念と合致する)。すなわち人格は意識によって構成され、意識とは記憶によって同一性を保証されたものである、とまとめることができる。

●ヒューム『人間本性論』

ヒュームは、個別の表象についてはその存在を認めるという点でデカルトを受け継いでいるが、しかしそれを統合し自己の同一性を保証するものとしての「観念」の存在は否定している。存在するものはあくまで個別の表象であり、その表象を感じている間のみ、私は私を捉えることが可能であるが、例えば深く眠っているときは表象を感じないため、ヒュームの論理上その間は「自己は存在していない」と考えることも可能である。同様に、自己の身体が崩壊したのちも、また同様に自己は無に帰すると考え、デカルト的、あるいはキリスト教的な「精神の不滅」を否定した。

私たちは、眠っている前後で自己が全く同一のものであることを信じて疑わないが、ヒュームによればそれすらも懐疑的なものとなる。これは、私たちが眠る前後で(周囲に関する)記憶の同一性を感じ取ることから、それを拡張して「自己の同一性」を信じているからである。個別の表象のみを信頼して、それを統合する観念の存在を否定するヒュームの立場からすれば、周囲の状況や経験という「対象の同一性」がたとえ保証されたとしても、それが「(認識主体である)自己の同一性」へ置き換わるプロセスは全く想像に由来するものなのである。

●カント『純粋理性批判』

カントはまず「直観の多様性」について言及する。あらゆる外発的な表象(経験的統覚)は、それに先立って与えられている純粋統覚に基づいている。つまり自己の具体的な経験的表象は、すべて「私の表象」、つまり抽象的で純粋な、直観の多様性を束ねて統一するものとしての私Iに基づいている。

(面倒なので英語で)例を挙げると、It is hot today.や、There is a desk.などの個別の表象(経験的統覚)は多様であるが、これらを統括するものとしての私Iが存在することは捨象されている。つまりこれらの文は本来、I think it is hot today.や、I think there is a desk.なのである。ここにおける従属節が経験的統覚であり、あらゆる経験的統覚に暗に付随しているI think が純粋統覚であるといえる。

カントはここから、(人間特有である)I think を持つ能力を「悟性」と称している。さらにカントは、精神異常に普遍的にみられる「常識の欠如」と「論理的強情」の性質から、我々が自己の悟性と他者の悟性を照らし合わせて「私的な思想」を律する営みを、「私的な思想の公的な表明」という社会的行為の中に見出した。

【2. 意識と無意識】

[概略]

前講で見てきた近代哲学における大きなテーゼは、「意識の統一が人格を支えている」という事であった。しかし、20 世紀にオーストリアの精神科医であるフロイトが「無意識」などの概念を提唱したことから、意識に関する見方は大きく変わっていく。本講ではフロイトの理論の変遷と、関連する記憶疾患について説明がされている。

フロイトの理論は 3 つの時期に区分することが出来る。第一期(~1897 年)においてフロイトは、開業医として多くの神経症患者への臨床を経験した。ヒステリー患者に対して、催眠療法を施した結果その症状が改善されることを発見したことから、神経症の症状を無意識の力の顕れであると考え、その原因を主に幼児期の性的な外傷体験(ex:性的虐待)に求めた。

しかし、1897 年以降フロイトは、そのような「外傷体験」の多くが、実体験ではなく患者の空想であることに気づき、これらの疾患の原因を外傷体験ではなく「内在的・本能的な衝動を組織化する」エディプス・コンプレックスなどに見出すようになった(第二期)。また、この内在的な、無意識の願望が、神経症のみならず意識一般の組織化の基礎となると考え、意識の三層構造を提唱した。

第一次大戦を経て、戦争体験による患者の臨床にあたったフロイトは、更に前述の三層構造に改良を加えた(第三期)。これら(第二期、第三期)は「精神分析」理論と呼ばれ、今日

まで続く学問体系を形成しているが、その基礎は「無意識」の仮定にある。この措定については、特に自然科学者などから多くの批判を受けたが、フロイトによれば意識の体系には必ず間隙が存在し、これを埋め合わせるものとして、忘却・抑圧などの要素を備えた無意識の仮定が有用である以上、科学的検証作業を積み重ねることによって無意識の存在を証明することができるかとされている。

講義内では、フロイトに関連して現代の記憶理論とその症状が紹介されている。外界の情報はまず感覚貯蔵にごくわずかな時間のみとどまり、その中で我々が注意を払った情報のみが短期記憶として保存される。この短期記憶を保つために心の中で何度もその情報を反復するリハーサルが行われて「記憶の固定」がなされ、感覚情報は長期記憶となる。これが記憶に関する現在の一般的なとらえ方であり、カナダの心理学者ハップによれば、これらは神経細胞の活性化、結合によって説明ができる。

この記憶に関する症状として、講義内では解離が何種類か取り扱われているが、その共通点として、外界からの強い衝動を原因とする、記憶の反復、あるいは抑圧が症状を引き起こすという点が挙げられるだろう。

[補足説明]

●第一期のフロイト

催眠状態の中で過去の苦痛記憶を思い出すことでヒステリー症状を改善させるという、催眠療法を発見した研究者はブロイエルである。フロイトはここから、過去の苦痛記憶、特に幼児期の性的な外傷体験は、「無意識な力」となると考え、また無意識の領域と意識の領域の「解離」、すなわち神経症症状は、この苦痛記憶からの防衛手段によるものであると考えた。催眠療法は、感情の放出によって「無意識な力」を解除する療法であるといえる。

着目すべきは、この時点では無意識と意識の解離が「症状」とみなされている点であり、その意味ではこの考えも、デカルト以来の「意識の統一」を正常とみなす概念の延長上にあるといえる。

●第二期のフロイト

ヒステリー症状の原因と見なされてきた「過去の外傷体験」の多くが虚構であることに気付いたフロイトは、その虚構を作り出す原因が人間の内在的・本能的な欲望にあると考えられるようになる。つまり、特殊な患者に限らず人間はみなその欲望を心の構造のひとつとして備えているのである。こうしてフロイトは、意識の三層構造(意識/前意識/無意識)を提唱し、内在的な無意識的願望の代表としてエディプス・コンプレックスなどを挙げている。

エディプス・コンプレックスとは、異性親への性的願望であり、この願望は社会的に善くないことであるために本来は抑圧されている。このように抑圧された願望が無意識の構成要素となっている。フロイトの精神分析療法は、これらの無意識的表象を、夢分析(前意識の分析方法の一種)などを用いて、意識的表象の領野にもたらずことで、症状の分析・改

善を試みる療法である。

●第三期のフロイト

意識の三層構造に改良を加え、エス/自我/超自我の三層を新たに提唱する。この理論の特徴は、三層がいずれも無意識の領野を含んでいることであり、特に人間の原初的な本能欲動であるエスや、親の存在の抽象化であり良心の基礎である超自我はそのほとんどが無意識であるとされる。エスの欲動や超自我による圧力、更に外的現実からの力を調停するのが自我である。

●PTSD(Post-Traumatic Stress Disorder)

心的外傷後ストレス障害。米国精神医学会が作成した PTSD の診断基準を参照すると、死への恐怖などを含む強い外傷的出来事の経験、その出来事を夢やフラッシュバックなどで再体験し続けること、外傷以前にはなかった感覚器官や日常生活、および心理動向の変化などが挙げられる。

たとえば東日本大震災で被災した一部の子どもたちは、しばしば「津波遊び」を行うことがあるが、これは子ども特有の「心的外傷体験の再体験」行為であるとされる。また、地下鉄サリンの被害者がその後一切鉄道車両へ乗れなくなった、などの例もある。

●解離性障害

解離性障害にはいくつか種類が存在するが、ここでも講義内の情報に絞ってまとめる。ファン・デル・コルクらによれば、解離はその発達段階に応じて三種類に分けることができる。一次解離は、自我と心的外傷の解離であり、PTSD がその代表例である。二次解離は、自我の不統合とも呼ばれ、離人症や痛覚麻痺など、体験する自我と観察する自我の解離が生じる。三次解離は、自我そのものが完全に解離し、解離性同一性障害などのように各人格が自己の中の別の人格を全く異なるものとして扱うことができている。

これらの解離の大きな特徴は、記憶の基礎となっている「自己」と「時間」の不具合である。解離はその段階が進むごとに自己の統括が困難になっていき、またその原因となっている心的外傷がいつまでも過去とならず心の中で反復・再体験し続けている。

【3. 意識と心】

[概略]

本講では主に初期の現象学、及び言語哲学が扱われている。

20世紀初頭に、オーストリアの心理学者であるブレンターノは「志向性」の概念を提唱した。ブレンターノはすべての心的現象を「志向的内在」によって説明し、この志向性の

概念はのちに現象学の理論的基礎となる。

その現象学を創始した哲学者が、同じオーストリアのフッサールである。フッサールの理論は、意識体験が常に何らかの対象を備えた志向性に基づく体験であるというテーゼに基づいており、それは「Consciousness is consciousness about something. (=Aboutness)」という言葉でも表される。またフッサールは『デカルト的省察』の中で、「コギトとコギタートゥム」や「地平」などの概念について言及している(後述)。

この「志向性」をめぐる議論について、講義内ではサールが多く引用されている。アメリカの言語行為論学者の大家であるサールは、心的状態の全てに志向性があるわけではないと考える。たとえば、信念や恐れ、希望などは志向的であるが、神経過敏、対象なき不安などは志向的ではない。志向的な心的状態はあくまでも全ての心的状態の一部にすぎない。更にサールは、志向的な心的状態をその「適合方向」に依りて三分化した。サールの理論はかなりややこしいので、まとめて後述する。

[補足説明]

●ブレンターノ『経験的立場からの心理学』

・志向的内在

心的表象は、必ずその表象とセットで「表象されるもの(=客体)」を含んでいるという理論。つまり、心的表象が志向するその対象は必ず同じ心の中に内在しているということである。例えば我々が外界の物理空間内の物体 A を観察するとき、心的表象として「A の知覚(表象)」が現れるが、この時外界の A とはまた別に、心理空間内に「知覚された A(=A')」が現れる。このように対象そのもの(A)と心的表象における対象(A')を区別したことが、ブレンターノの大きな功績であるとされる。

この理論は、たとえば本当は見えないはずのものが見えるという錯覚の現象は、A は存在しないが A' が表象されるからである、というように素朴に應用することもできる。つまり、デカルト的な「意識がまず存在し、それが対象を確認する」のではなく、意識と対象が相関的であることを指摘している。

●フッサール『論理学研究』、『デカルト的省察』

・志向性

ブレンターノに師事したフッサールは、彼の理論を継承して志向性概念を発展させた。たとえば対象 X について、知覚主体 A、B、C が各々まったく独立しているにも関わらず「同じ対象 X を見ている」と認識しうる理由について、フッサールは「志向的本質の同一的統一性」のみが重要であるため、と述べている。つまり、体験の記述的内実の相違(A、B、C)を考慮に入れながら、それらが「同じ対象 X」を志向しているという志向的本質こそが、「同じ対象を知覚している」という統一性を保っているということである。

・コギトとコギタートゥム

ラテン語のコギト cogito は「I think」の意味であり、デカルトのテーゼにもあるように思考する主体を指す言葉である。一方のコギタートゥム cogitatum は cogito の受動態で、「what is thought」、つまり思考される対象を指す。コギトは常にコギタートゥムを伴っている。つまりコギトとコギタートゥムの関係は、ブレンターノが提唱した志向的内在性の概念とほぼ同じであり、両者が相関することの重要性を明らかにするためにフッサールが取り上げた概念であるといえる。

・地平

たとえば我々がタバコの箱の表面を知覚するとき、箱の裏側の見えない箇所についても、これまでの経験などからその形状などを暗に知覚している。ここで、箱の表面という顕在的側面に対して、見えない裏面のことを「地平 Horizont(潜在的側面)」とフッサールは呼ぶ。

地平とは我々の意識全般を構成するものであり、それは過去の知覚の蓄積(タバコの箱には裏側もあるものだ)という「歴史的地平」と、現在の状況の知覚(タバコの表面が見えているのだから裏側もあるはずだ)という「現地的地平」の融合 Horizontverschmelzung によって構成される。歴史的地平は通常は経験の中で類型化され、無意識の中に「沈殿」していき、潜在的な志向性として、タバコの箱と向かい合った時に発揮される。これによってタバコの箱の裏側は、未だ知覚されていないが「知覚されうるもの」として思念されるのである。

このように、フッサールの志向性に関する議論は、実際に現前するもの以上のものを人間が思念するという特質、つまり「人間の意識が自らを超えて思念する」という意識の大きな特徴を捉えようとしたものであると総括できる。

●サール『志向性』

・適合方向

前述したようにサールは、心的状態の中で志向性を持つものを、適合方向に応じて三分化した。ひとつは、「mind-to-world」的適用である。たとえば「ああ、今雨が降っているな」という「信念」がそれである。これは心 mind が世界 world に従属している、つまり心が世界へ適合している(心→世界)状態のことである。

二つ目は「world-to-mind」的適用である。「雨が降ってほしい」という「願望」がそれである。一つ目と反対に、願望の成就のためには世界が心へ適合する(世界→心)必要があるからだ。

三つ目は「適合方向ゼロ」と講義内では紹介されていたが、mind と world の関係ではややわかりにくいので、言語行為論的に、言葉 word と世界 world の関係から適合方向を捉えなおす。まず word-to-world の適合方向においては、言葉は世界に適合しているために、言葉は世界を正しく描写する必要がある。先ほども例に挙げた「いま雨が降っている」などの信念表明などがそれである。反対に world-to-word の適合方向においては、

まず言葉が優先し、それに世界が適合するような言語運用を行う。「傘を貸してほしい」などの、命令や約束、願望がこれにあたる。

そして三つ目は方向ベクトルがゼロ、すなわち word と world が既に一致している状態を指す。つまり世界と言葉が既に一致していることを前提とする、謝罪や感謝などがそれである。例えば「傘を貸してくれてありがとう」という言葉において、「傘を貸す」という word は既に、傘を貸してくれたという事実 world と一致していなければ言葉として成立しない。

・充足条件

志向的であり、かつ適合方向を持つ状態については、必ず充足条件が存在する。充足条件とは、その志向的状态が充足しているならば当然成立しているだろう条件のことである。例えば「いま雨が降っている」という信念、また「雨が降ってほしい」という願望を充足させる条件はいずれも「雨が降っている」ということである。

・志向性と意識/因果

サールによれば、志向性は必ずしも意識の上に存在している訳ではない。例えば、自分の10世代前が人間であるという信念、また「生きたい」という願望が常に意識の俎上に載っているわけではない(かといって全く存在しないわけでもない)。他方で、視覚などの知覚経験は必ず「意識されている心的状態」である。ここに、志向的状态を、その因果から二種類に分けることができる。

一つは、視覚経験などの意識的な志向的状态である。この時、原因は世界にある対象との出会いであり、そこから結果として視覚経験という志向的状态が生まれる。もう一つはその反対で、あらゆる行為のように、私たちの志向的状态を原因としてその結果として何らかの行動を行うという状況である。このように志向的状态は、原因とも結果ともなりうるのである。

●アスペクト変換

「うさぎあひる」や「ルービンの壺」のように、複数の相が提示された場合、我々がその内のどれを志向的对象と見なすかによって像に変化が生じる。これは、我々が普段いかに無意識のうちに志向性を承認しているかを示す好例である。

●クオリア qualia

現象的な質 phenomenal quality の略語とされている。例えば我々が波長 630～760nm の際立った光を受容するときに「これは赤だ」と感じるときの感覚がクオリアである。そもそも心的概念は「現象的」なものと「心理学的」なものに分けることができ、前者がクオリアであり、後者はサールが問題としたような志向性を備えた因果を含む心的状態である。

心理学的な心的概念は、因果関係を基礎に考えるものであり、オーストラリアの哲学者

チャーマーズによれば、それは easy problem である。例えば「怒り」という心的状態と「殴る」という行為は、心理学的に見ればともに物理システムとして規定することができ、二者の間の因果関係を解明すれば事足りる。一方で現象的な心的概念を問題にする場合、たとえば「怒り」とはどのように感じるものなのか、といったクオリア(質)を解明するのは容易な問題ではなく、これが hard problem と呼ばれる。

【4. 意識と脳】

[概略]

20世紀前半、ワトソンやスキナーによって行動主義の心理学が興された。行動主義とは、客観観察可能な行動のみから科学的にアプローチを行う心理学であり、行動を主眼に置き意識は対象とされていなかった。

20世紀中盤に入って、プレイスとスマートによって、心的プロセスは内的で個人的なものであり、客観的な観察は不可能であるという主張がなされた。彼らはこの他にも、心身二元論を否定したうえで、心的プロセス(意識)は脳のプロセスであるとして「心脳同一説」を提唱した。プレイス・スマートは心のプロセスを意図と感覚とに分け、前者は行動主義的に説明可能であっても、クオリアを含む後者は説明できないと考えた。この説を発展させ、二者をともに脳の物理化学的状态のみで説明可能であるとするアームストロングのような、更に進んだ心脳同一論者も現れた。また、心脳同一説の別なアプローチとして、脳の非物理的属性から心的状態を究明しようとする、パトナムらの「機能主義」も勃興した。なおこの動きに対し、サールは「生物学的自然主義」を採用して対抗し、心的現象は植物の光合成などと同様に生物学的な現象であると捉えた(ややこしいので後述)。他にも心的現象を説明する学派として講義内では、チャーチランドによる「消去主義」が紹介されている。

さてパトナムは、「水槽の中の脳」という思考実験を提示して、そのうえで「我々が水槽の中の脳ではない」ことを著書の中で証明しようとしている。「水槽の中の脳」とは、我々が実は邪悪な科学者によって脳を摘出されており、その脳が培養液の中でコンピュータと神経接続されて、コンピュータによって諸々の経験を脳に送り込まれてそれを「経験」と感じているに過ぎない、つまり我々は実際には外界を全く経験していない可能性があるという仮説である。パトナムはこの命題の自己論駁性から、命題の偽を証明するが、我々が「水槽の中の脳」仮説が現実にもありうるかもしれないと思ってしまう理由を2つ挙げている。ひとつは、「水槽の中の脳」が物理的に可能な世界であるという想起から、それが現実として存在しうることを素朴に信用してしまうこと。もう一つは「指示の魔術説」である(後述)。

[補足説明]

●機能主義 Functionalism

パトナムによって提唱された説で、心脳同一説の亜種。スマートらが心的プロセスを脳の物理的属性と同一視していたことに対し、パトナムは非物理的な属性を重視する。たとえば、コンピュータには形状や質量、どのような部品があるかといった物理的属性のほか、プログラムなどの機能的属性が存在する。機能的属性は、コンピュータの物理的存在とは全く無関係に実現しうるため、パトナムはこれを非物理的な属性と見なす。同様に脳についても、その機能的な側面を重視するのが機能主義の立場である。

機能主義においてはその機能を重視するために、「怒り」という心的状態から「殴る」という行為に及ぶことと、そのような脳の状態から「殴る」という行為に及ぶことの因果が同一であることから、「怒り」という心的状態とそのような脳の状態を「結果として同一」と見なす。ここから機能主義の強みは、このシステムが人間以外にも適用可能であるという点であることがわかる。つまり機械でもこのような機能を脳と同等に実現しうるということだ。しかし、パトナム自身が指摘しているように、機能主義は感覚の質的な差異、すなわちクオリアの存在を捨象しているために多くの批判を受けている。

●生物学的自然主義 Biological Naturalism

サールが機能主義に対する批判として提示した説。心的状態は「脳の作用を原因とする (caused by)」上に「脳の構造・システムにおいて実現している (realized in)」と説明される。サールはこれをわかりやすくするために水を例に挙げている。

水の液体としての諸性質 (=心的状態) は、水分子の行動 (=脳の作用) を原因としており、またその諸性質は水分子の集合 (=脳の構造・システム) において実現されている。液体 (のシステム) はあくまで水分子のみで構成されているにも関わらず、その分子もまたシステムの流動性の影響を受けている。端的に言えば「相互作用の」関係になっているということである。つまり、心的状態はあくまで脳の作用によって生じてはいるものの、脳の機能による純粋な産物というよりもむしろ相互作用の中で互いに成立しているものなのである (たぶん)。

たとえば視覚経験は脳の働きを原因とし、また視覚経験は脳の構造において実現されている。同じように、「渴き」は視床下部に至る神経作用を原因とし、視床下部において実現している。

●「水槽の中の脳」の否定

「水槽の中の脳」のテーゼは、我々は実際には外界を全く経験しておらず、すべては作り出されたイメージに過ぎないということである。これを仮に真とすると、その偽のイメージの中には当然、我々が今問題としている「水槽の中の脳」のイメージも含まれる。すると命題より、我々が想像しうるイメージはすべて偽であるので、この「水槽の中の脳」

というイメージもまた偽となる。すると、「我々は水槽の中の脳である」という命題も偽となってしまうので、命題は成り立たない。

このように命題の中に反証が内在しているものを自己論駁的といい、「私は嘘つきである」のような自己言及のパラドックスがその代表例である。

●指示の魔術説

例えばアリの足跡がチャーチルの顔のように見えるのだとしても、それはアリが意図的にやったことではなく、つまり図は必ずしも指示性・対象を持っているわけではない。

逆に、指示性が成立するためには、外在的・物理的な対象 A と、それを表象する心的な像 B の両方が不可欠である(B→A という形で指示する)。しかし我々は時として、心的表象 B しか持っていないにも関わらず、あたかも指示が成立しているかのように考えてしまうことがある。これが指示の魔術である。たとえば、「水槽の中の脳」において心的表象 B はコンピュータによって成立しているが、それと対になるはずの外在対象 A は存在していない。これでは指示が成り立たないので、人間本来の志向性自体がないことになってしまう。

つまり、単にイメージを抱くこと(B)と、指示すること(B→A)は異なるということである。

●消去主義

素朴心理学においては、たとえば「顔の紅潮(A:観察可能)」→「怒っている(B:観察不可能)」→「殴る(C:観察可能)」という他者の一連の動きを、観察不可能な領域も含めて予測し、記述することが可能である。

チャーチランドによれば、この流れの全てはいつの日か神経科学の用語によって置き換えが可能であり、神経科学が完成すれば、これまで我々が心の理解に用いていた諸々の用語・概念はすべて「消去」され、神経科学の概念によってまったく新しく再構成されるだろうとされている。これによって心をめぐる諸理論は、物理科学一般とより強力で統合することが可能であるとチャーチランドは主張する。

なお、ここでいう「消去」とは概念や言葉が積極的に抹殺されることではない。そもそも概念や言葉は実体として存在するものではなく、人間の理解の推移に応じて絶えず変遷され、淘汰されていく。この動きの中で、たとえば中世においては広く信じられていた「魔術」の存在が、自然科学によって説明されるのではなく、言説から自然と消去されていったように、心的状態という概念そのものも消去されていくだろうとしている。

【5. 意識と物】

[概略]

本講では、人工知能探求の歴史が説明されている(長い)。

20世紀前半の計算機科学者チューリングによれば、デジタル計算機は万能機械であるがゆえに、処理速度の差異を除けばそれらはすべて同等であるため、コンピュータを考えるとときにはそのようなハードウェアでなく、プログラムなどのソフトウェアがその本質であるという考えを提唱した。そしてチューリングは「チューリング・テスト」と呼ばれる、機械が知的かどうかを判定するためのテストを考案した。テストに合格すれば、その機械は人間と同等の知能を持った「人工知能」と見なされる。

これに対し、チューリング・テストに否定的であるサールは「中国語の部屋」と呼ばれる、チューリング・テストのパロディのような思考実験を提示して、テストの弱点を衝いた。サールは、コンピュータが人間の心をシミュレートするという「弱い AI」論は認めつつも、人間の心そのものをプログラムと見なして機械で全面的に再現可能であると考え「強い AI」論には否定的であった。

他方で、「強い AI」論者である現代の発明家カーツワイルは、チューリングの説を発展させて、心においてもハードウェアは脳と神経以外の、優れた電子回路で置き換え可能であり、その根拠を脳細胞のパターンの分析に求めた。つまり、アイデンティティや意識は脳細胞のパターンによって構成されたものであり、これを解明すればハードウェアは置き換え可能であると考え、スピリチュアル・マシンの理論を提唱した。

これとは反対に、アメリカの哲学者ドレイファスは、コンピュータが人間と同種の物質的身体を所有していない以上、人間のような知的振る舞いは実現しえないと考え、チューリングら人工知能を肯定する人々が、プラトン・デカルト以来の古典的な精神の優越に基づく二元論に囚われているとする「反 AI 論」を展開している。

人工知能の研究過程ではさまざまな問題が提起されている。例として、外界の無限の状況に対して情報処理をどの程度の枠組みで遂行させればよいのかを問題にした、マッカーシーらの「フレーム問題」や、入力された情報の中からどの部分を抽出させるかという判断を論点とした、橋田浩一らの「部分性」問題などが挙げられる(後述)。

これらの議論では知的システムにおける身体性が着目され、その動きの中で MIT のロボット研究者であるブルックスは、「表象なき知性」という概念を提示して、従来の AI のような知識(表象)のインプットから、環境との相互作用(インタラクション)の中で知性を創発させようとする理論を打ち出した。

今日の人工知能研究については、日本と欧米では動向が異なるといわれている。MIT をい始めとする欧米においては全身型のヒューマノイドを作ろうという動きは、宗教的倫理から忌避される傾向にあり、それゆえ人体の機能の一部(手足や顔の表情など)を実現させようとする方向へ努力が注がれている。他方、日本ではそのような傾向は少なく、欧米の研究者が日本で研究を行う例もあるという。この動きの中で日本では、1970 年代と早い段階から全身型ヒューマノイドの研究が進んでいる(ex:早稲田大の WABOT、ホンダの ASIMO など)。それ故に、ロボットが実際の人間に近づいていく過程で生じる「不気味の谷」の問題についても、日本では実用的観点から多く議論される傾向があった(後述)。

講義ではさらに、人工知能の研究について内部構造のメカニズム解明よりも求められる機能の実現に(ひとまず)重きを置く「構成論的アプローチ」の考えについても触れられている。このアプローチの日本における代表例が、大阪大の石黒浩教授であり、彼は表面の模倣に特化したロボット研究を推進している。石黒教授のようなロボティクスの研究分野と、認知発達研究の分野が互いに影響しあう動きが現在見られてきているが、その基礎となっているのが「構成論的アプローチ」である。ロボティクスの分野において構成論的アプローチによって「人間の心という機能を持ったロボット」が実現されれば、それを徹底的に分析することで(=介入研究)人間の心がわかる。しかし、分解しても心に相当するものは存在しない(と石黒教授は言う)、すなわち「機能」としての心は存在しない。しかし、(石黒教授を含めて)我々は心の存在を「実感」している。石黒教授によれば、これこそが「人間」であることの証であり、この実感を持たない者が機械なのである。

[補足説明]

●チューリング・テスト Turing test

「人工知能は人間と同等に思考することができるか」を測定するために開発されたチューリング・テストの仕組みを以下に記す。実験者 A は、別の部屋に存在する B と文字を通じて会話をを行い、また別の部屋に存在する C とも同様に会話をを行う。ここで、実験者 A が B(人間)と C(実は機械)を確実に区別できなかった場合、機械 C は人間と同等のコミュニケーションが出来るだけの知能を備えていると見なされる。チューリング・テストがこのように人工知能の機能的側面のみを重視していることから、このテストが機能主義に基づくものであることがわかる。

なお、講義内では触れられなかったが、このテストにおいて実験者 A が、B と C のうちいずれかが機械であることをあらかじめ認識しているかどうかは、チューリング自身は言明していない。しかしこの前提条件がテストの結果に大きく響くことは、多くの学者に指摘されている。

●中国語の部屋 Chinese room

機能主義に基づくチューリング・テストを批判するためにサールが考案した思考実験であり、実験そのものの欺瞞を暴くために実験の構図がチューリング・テストと類似している。「中国語の部屋」と呼ばれる部屋に、英語しか読み書きのできない英国人が閉じ込められている。部屋の外からは中国語の文字の羅列が提示され、どのような文字列に対してどのような文字列の返答を行えばよいのかがすべて記載されている英語の手引書が部屋の中に存在する。英国人はこの手引書通りに返答を行えば「中国語の部屋」の住人としての役割を果たせるし、外から見ればあたかも「中国語の部屋」は中国語を理解しているように認識される。

サールは「(アルファベット使用者にとって)意味不明な文字列」という意味で中国語を引

き合いに出しているの、そこは注意すべきであろう。ともあれ、このたとえの中の中国語が人間のコミュニケーション言語であり、英語がプログラム、部屋が人工知能の筐体であり、部屋の外から文字列を提示する人が、チューリング・テストにおける実験者 A であることは明白である。

この思考実験からわかるサールの見解は、人工知能が「弱い AI」的能力しか持っていない、つまり本当に心を理解しているのではなく、それをマニュアル通りにシミュレートしているだけということである。つまり、言語哲学的に換言すれば、コンピュータは統語論的な操作しかできないのであり、意味論を兼ね備えていない。どれほど完璧な文法に基づく文をコンピュータが提示できるのだとしても、それはコンピュータにとって何の意味も持っていないということなのである。また、生物学的自然主義の立場をとるサールは、そもそも脳という特殊なハードウェアを持たない人工知能はその時点で、脳が存在する場合と同等程度の心を作り出すことができないと考える。しかし逆に、サールによれば、より脳に近い有機的なものを人工的に作り出すことが可能である場合、脳が作り出すそれに類似した心というものを作り出すことは否定できないという。

●フレーム問題

ロボットが外界からの情報を取り入れる際は、プログラムへの入力ののちに表象へ翻訳されるプロセスが必要であるとされる。しかし、たとえばロボットがあるひとつの動作をする場合にも、外界には無限の情報が存在するのであり、そのうちどれを「動作の遂行のために必要である」と見なして処理させるか、つまりどのように情報処理の枠組みを設定するかという問題が「フレーム問題」である。換言すると、世界のある状態に対して、ロボットが何らかの働きかけをすることによって予測される「世界のほかの状態」への移行を、どの程度書き出す必要があるのか、という問題である。

この問題はつまり、人間でいうところの「常識」をどのようにしてロボットに備え付けるかであり、ひとつの対処方法としては膨大な経験のデータベースを蓄積して「常識」に類似する行動原理の源を作り出すというものがあるが、それでも外界の情報は無限であるしデータベースもそれに応じて無限に増えていく。そもそも、人間自体いったいどうやってこの「フレーム問題」を切り抜けているのか解明されていないので、この問題は未解決であるといえる。実際、人工知能の世界においてはこの問題を回避するために、ボードゲームなど外界からの情報が有限個である場合に的を絞って開発がなされている状況も存在する。

●情報と部分性

フレーム問題と類似している。そもそも AI 研究がこれまで発展してきた理由は、情報のうちに存在する機能的側面のみに着目して、その物理学的詳細(=媒体)を捨象してきたからである。つまり情報を処理するときに、そのモノ性、身体性を捨象していたのである。

しかし実際には、たとえば黒板に書かれた板書をメモするとき、我々はその内容に着目しているものであり、独特の筆跡などまで逐一模倣することはないが、単語間の位置関係など理解に必要な情報には注意を向ける。ここで我々はすなわち、情報の中で一部分のみを切り出して処理しているものであり、どの情報进行处理するかという選択においては、やはり「常識」に基づいた判断がなされている。情報の有用性は、世界からどの部分を切り出すのかという選択行為の中に存在するのであり、その中には機能的側面(板書の内容をどこまでメモするか)のみならず、身体性(板書の様式・形態をどこまでメモするか)も含まれている、ということであり、AIはそのうち前者のみに目を向けてしまっている(からこそここまで発展してきた)というのが「部分性」の議論である。

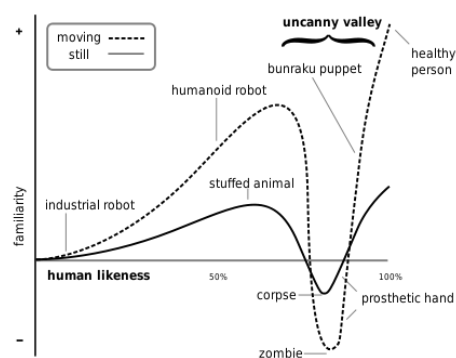
●表象なき知性 Intelligence without representation

古典的なAI研究においては、外界の認識に際してはその都度プログラムに入力して、それを表象として再構成するプロセスを経ていた。たとえばロボットがドアノブを回すときは、「ドアノブを開ける」という(言語的)表象にロボットがアクセスする必要があり、その時ドアノブに何らかのイレギュラーが認められる場合、その都度ロボットは表象を構成しなおす必要がある。この煩雑な過程の中で、どこまでをプログラム処理するか、どこまでを表象として認めるかといったフレーム問題などの障壁が生じている。

ブルックスはこのプロセスから脱却し、表象への再出力を介在させずに、ロボットが外環境と直接触れ合うことで、つまり外界を実際に動き回る過程で、表象なしに知性を発達させていくAI開発を提唱した。別な言い方をすると、これまで情報の抽象化、つまり取捨選択は研究者のお膳立ての上で行われていた。しかし、真にロボットが「知的」行動をとるのであれば、その情報の抽象化をこそ自らの知性で執り行うべきである。この理念のもとでブルックスが開発しているものは「行動規範型制御アーキテクチャ (BBCA: behavior-based control architecture)」と呼ばれているらしいが、難しすぎて私にはもうよくわからない。

●不気味の谷 Uncanny valley

日本のロボット工学者である森政弘の提唱した現象(右図:Wikipedia より引用)。横軸はロボットの人間への類似度であり、縦軸はそのロボットに対する我々人間の感情的反応である。当初の段階では、ロボットが人間の特征に近づくにつれて、我々がロボットへ感じる親近感は比例して大きく



なっていく(工業用ロボットから通常のヒューマノイドへの推移)。しかしある段階から、人間は急激にロボットに対して嫌悪感・不快感を抱くようになる。その「不気味の谷」を通り過ぎると、親近感はV字谷のごとく回復し、そしてグラフは類似度100%である通常の

人間に至る。

この現象の原因としては、人間がある段階まではロボットの自分たちとの類似性のほうに着目するが、ある段階からは類似性が大きくなったがゆえにロボットの「非人間的部分」、つまり自分たちと異なっている部分に目が行きがちになるから、という説が考えられる。特に、細かい外観の特徴や動作の不穏さなどは、ゾンビや病人(とりわけ精神障害など)を想起させるがゆえに、否定的な感情が先行するという説もある。

グラフィックス技術を駆使した映像作品などで、コンピュータグラフィックスで作成された人体に対して「不気味である」という感覚が生じる原因もこれと類似したものであり、そのため映像作品ではあえてリアリティを追求せず不気味の谷の前段階に踏みとどまってデフォルメを施すことで、観客の親近感を保持しようとするものもある(ex: Mr.インクレディブル)。なお、森自身も同様に、ロボット開発においては不気味の谷の前段階で踏みとどまるべきだという考えである。

●構成論的アプローチ A Constructive Approach to Intelligent Robots

人間の認知メカニズムの解明を優先するのではなく、むしろ「人間のような機能を持ったロボットを作る」ことを以て理解に替えようとするアプローチ方法。つまり、人間としての機能が実現されれば、その内部構造の詳細が解明されなくてもよいとする考え方。そのアプローチ方法ゆえに、ロボット工学のみならず認知発達科学の分野との横断研究もみられる。

ブルックスが批判したように、古典的 AI 研究における人間知性の理解は「人間の行動原理である膨大な『常識』を、プログラムとして書き出すことによって実現する」思想に基づいており、構成論的アプローチはこれに対抗する思想となっている。

【6. 心と身体】

[概略]

前講の後半ではロボットにおける身体性・環境との相互作用について述べられていたが、本講ではこれに引き続き、人間の認知における身体性に関して議論がなされている。この議論の一つの中心となっている考えが「身体性認知」と「拡張された心」であり、二者のはいずれも、身体や外環境との関連の中で人間が認知を達成していることを示している。

ポルトガル出身の神経科学者ダマシオは、「ソマティック・マーカー」仮説を提唱し、人間の認知機制がむしろ脳の情報処理に先行して身体による介入がなされていることを主張した。他にもダマシオは、神経的表象やセルフなどの概念について述べているが、これらについてはまとめて後述する。因みにダマシオは、「水槽の中の脳」についてパトナム同様に反証を提示しているが、その根拠は「身体なしに脳は心を成立させられない」という考

えに基づいている。端的に言えばダマシオは、従来の神経科学の主流であった心をすべて脳に還元する手法に反対して、身体それ自体の重要性を説いているのであり、むしろ人間の身体があったからこそ、脳が現在のようにまで進化しえたのであり、それが心を生み出したのであると述べている。

通常の進化論から考えれば、運動反応が発達することが有用であるとしても、心的反応の発達には不要に思える。たとえば、ハチが向かってきたときにすぐさま逃げるといった反応は重要であるが、それに伴う「怖い」、あるいはより原初的な「ヤバい」というとっさの心的反応は本当に必要であるのか。ダマシオはこれについて、むしろそのような心的反応があるからこそ、優れた近くに基づく対象へのより正確な認識や、運動反応の改善、そしてモニタリングによる将来の予測などが可能になるとしている。すなわち心が生まれた原因は、身体や外環境のイメージを体内で再構築することが、身体や反応にとってより有利であるが故の進化なのである。人間が外的環境に対処するために進化させてきたのが「心」なのであり、その意味では他の動物たちもそれぞれ、外的環境に備えた「心」のようなものを所有しているといえるだろう。

しかし、心に関連した人間特有の能力として、他者が自分と同じような心を持った者として互いにそれを理解しあえることが挙げられる(cf:「心の理論」)。人間がこのような能力を相互に身につけることで、社会的やり取りの性格は他の生物と大きく異なるものへと発展していき、さらにそれらは世代を経て更新され、物質的・記号的人工物(石版から口承まで)を媒介として蓄積されていった。人間の認知発達を考える上では、3側面に基づく認知発達の研究が求められているとされる(後述)。

[補足説明]

●身体性認知 Embodied cognition

アメリカの心理哲学者シャピーロなどによって近年提唱されている考え。シャピーロによれば、我々がどのように世界を認識するかという認知の方法は、我々がどのような身体構造を有しているかという物理的制約に大きく依存している。また、認知の核にあるものは表象的プロセスではなく、むしろ環境と相互作用を行う有機体の身体なのである。アルゴリズムに頼った認知過程という古典的AIの仮説から、環境との相互作用を重視する仮説への移行は、ブルックスやダマシオらが主張していることと同様である。

身体性認知の仮説で最も重要なことは、身体および世界(=外的環境)は、認知の因果要素であるのみならず、認知の構成要素でもあるという主張である。たとえばお腹がすくとイライラしてくるといった認知行動において、「お腹がすく」という身体性は、「イライラする」という認知に従属する単なる外的原因なのではなく、それ自体が認知を内的に構成するひとつの骨組みなのである。

●拡張された心 Extended mind

身体性認知同様に、近年多くの哲学者・科学者らによって提唱されている考え。ウェールズ出身の哲学者ローランズによれば、心的プロセスは4つのeに基づいているとされる。embodied が身体性認知であり、embedded と enacted はブルックスらの主張する環境との相互作用という枠組みで心的プロセスを考えるものである。extended とは、心的プロセスが脳・神経学的プロセスという体内の枠組みから次第に外界を巻き込んで広がっていく様を示している。これが「拡張された心」の考えである。

たとえば、コンピュータに表示されたある図形があるソケット(溝)にすっぽり収まるかどうかを解く図形問題を考える。図形を心の中で回転させるか、指でボタンを押して図形を回転させて解くか、脳にインプラントを埋め込んでコンピュータを操作して回転させるか、という3つの手法には実は何の差もない。すなわち我々は、言語や文字、自然物などの外部メディアを利用・加工することで、それらを認知プロセスの補助として用いているのである。つまり、認知プロセスのうちのすべてが脳内で完結する必要はないということだ。

それゆえたとえば、Aは美術館に行きたいと考えたときに脳内で「美術館が53番通りにあること」を思い出して、53番通りへ向かう。一方、記憶障害を持つために常に(普通ならば脳内にあるはずの)情報を参照するためのノートを持ち歩いているBは、同じく美術館に行きたいと考えたときに、ノートを開いて「美術館が53番通りにあること」を発見して、53番通りへ向かう。二者において、「美術館が53番通りにあること」という信念は同じであり、その信念を参照するための外部メディアが脳神経かノートだけの違いであり、「拡張された心」の理念に基づけばAとBの認知プロセスは同質である。

●ダマシオ『生存する脳』

・ソマティック・マーカー Somatic marker

身体標識。不穏な対象に対して、推論やコスト・ベネフィット分析に先立って反応を始める「不穏な直感」、すなわちネガティブな状況について我々の注意がほとんど自動的に及ぶような身体的反応に関する仮説である。この自動化された反応によって、我々がリスクに対して取るべき選択肢は大きく縮減していき、(今度は推論やコスト分析などに基づいた)行動の選択が非常に容易となる。ダマシオはこの仮説を立てたうえで、心と身体が分離しているのではなく一体となって認知プロセスを構築しているとする自説を補強した。

・神経的表象 neural representation

ダマシオによれば、すべての有機体が心を持っているのではなく、有機体がイメージとして、あるいは思考プロセスを補強する道具として、あるいはそれに基づく将来の予測や行動に影響を与える要素として機能する「神経的表象」を持つ有機体のみが、「心を持つ」といえる。神経的表象はニューロン回路の生物学的変化によって生じるものである。

・セルフ neural self

ダマシオによれば、神経的表象によって出来た我々の生物学的状態は常に再構築を続けているのであり、これを維持するためには脳システムと身体システムが全面的に機能し続

けていなければならない。たとえ一時的であっても、部分的であっても、これらのシステムに何らかの支障が生じた場合、我々の神経的表象に何らかの変化が生じ、そこから我々の心の状態は著しく変化してしまうだろうとしている。

●文化的認知

霊長類の研究などにも携わっている、ドイツの進化人類学研究所の所長でもある認知心理学者トマセロによれば、人間の認知発達を理解するには以下の3つの時間枠に着目する必要がある。

・種としての発達

霊長類の中のヒトという種が、同種の他の個体をどのように理解するようになっていったのかという、系統発生的(生物学の一派としての人類学的)な発達の時間枠。

・外部メディアへの発達

人間が、「種としての発達」のもとで構築してきた独自の他者理解(社会的理解)の手法を、物質的・記号的な人工物などを經由して継承・更新・蓄積していった方法とその推移に関する、歴史的な時間枠。

・個としての発達

人間の子どもが、自分の属する社会の文化をどのように受容し、どのように認知プロセスを発達させていくかといった、発達心理学的な、個体発生的な時間枠。

【7. 意識と世界】

[概略]

この講はおそらく、次講の前提知識である近代から現代にかけての哲学における他者・世界認識について、基礎的なことをおさらいしただけと思われる。ので、詳しくは補足説明で哲学者ごとに分けて書く。

[補足説明]

●カント『純粋理性批判』

カントは人間認識についての問題を考える際にまず、「主観的認識の客観化がなぜ可能なのか？」について言及する。カントによれば、主観的認識が客観化されるのではなく、主観自体が客観的な制約に従うがゆえに、(アプリアリな)主観は客観的な世界を認識できるのだと考えた。これをコペルニクスの転回という。無限に広がる客観的世界から人間の認識について考察することは事実上不可能であり、そのため逆に制約された人間の認識が客観的世界を形作っているのだと考えることで、世界を把握しやすいものとしたカントの転回法を指す。

このため、前提となる人間意識がきちんと統一されていることが、客観的認識を可能とする条件となる。これは超越論的統覚によって司られている。あとは第 1 講で説明されたことと同じである。

●フッサール『イデー』

フッサールによれば、我々が普段経験している現実世界は、存在定立(それが存在しているという信念)によって成り立っている。そこでは諸々の知覚によって、それが存在しているという「意味」となって我々にそれが与えられている(すなわちフッサールによれば存在と意味は別物である)。この現実世界を無批判に受け入れている状態を自然的態度という。

ここで、我々の存在定立を支える信念をすべてエポケー(カッコに入れる／判断停止を行う)すると、個々の知覚も、人間という自我もエポケーされる(エポケーは消え去るという意味ではないので注意)。すると何も残らないのかというと、ここで純粹意識と呼ばれるものが立ち現われる。純粹意識はこの意味で、デカルトのコギトと類似しているが、デカルトがコギトを実在するものと考えたことに対し、フッサールはあくまでも人間経験を可能にする条件・原理として純粹意識を提示している。

●ウィトゲンシュタイン『論理哲学論考』

ウィトゲンシュタインによれば、世界とは事実によって構成された総体であり、事実は論理によって写しかえられる。そしてその事実の論理像こそが思考であり、言語であると考えられる。すなわち思考・言語の総体こそが世界の像をなしている。このように考えると、「5.6 私の言語の限界が私の世界の限界を意味」し、その外から私は出られないということが云える。

すなわちウィトゲンシュタインは私そのものが世界であるという論理を支持し、それゆえに独我論を肯定する。なぜならば世界の外というものは存在せず、世界の中に主体は存在しえないからである。たとえば眼と視野の関係について、眼は視野の中には存在しえないが、世界の外という論理空間は存在しないため、当然の帰結として「眼そのものが世界である」と云える。つまり、自我は世界内存在ではなく世界そのものであり、哲学的自我は主体の世界の限界を示しているのである。

●サルトル『存在と無』

「まなざしの相克」について。他者とは私をまなざす(対象化する)存在であり、私が誰かをまなざす(対象化する・志向する)ときは必ず相手も私をまなざしているのである。このまなざす・まなざされるの相克関係が、自己と他者の関係を規定するものであるとサルトルは考えた。講義内ではこれに関連して、フォーコーがベンサムを引用して提唱した、概念としての「パノプティコン」「逆パノプティコン」などが紹介されている。

【8. 自己の意識と他者の意識】

[概略]

本講では、自己の意識と他者の意識の差異、及び共感などについてどのように説明するべきなのか、様々な側面から検討がなされている(長い)。

講義では最初にニーチェの思想が引用される。ニーチェは、我々がいま存在している「現象の世界」のほかに「真なる世界」があるという思考は、生への疲労からくるものであり、実際にそのようなものはないと考えた。同様に、目的と手段、主観と客観などの関係もそれらは解釈に過ぎず、それがそのまま事実ではない。では必然的なものとは何かというと、ニーチェによればこれらはすべて「力への意志」から発生しているという意味において必然的なのである。認識や真理の希求もすべて、力への意志から派生したものだとした。このように、認識に絶対性・客観性を認めず、すべてはパースペクティブの差異によった解釈であるとするニーチェの考えを「パースペクティブ(遠近法)主義」という。

ニーチェの「解釈」の考え方に関連して、『解釈学』を著したドイツの哲学者であるディルタイが紹介される。ディルタイによれば、自然科学の基礎づけは世界の数量化であることに對し、精神科学の基礎づけは「体験・表現・理解」にあるとした。A の内的なものが表現され、それが B に理解あるいは解釈されることで拡張する。このように理解・解釈を通じて個人の内的なものは拡張され、やがて精神科学一般を規定するものへと至るとディルタイは考えた。

以降では、分析哲学者ネーゲルからデネット、ハンフリーに至る「クオリア」問題の推移が概観されている。クオリアの概要については第 3 講で述べたとおりであるが、クオリアがそもそも存在するのか、だれが持つのか、なぜあるのかなどの議論は脳科学や心の哲学の一大問題(チャーマーズの云う hard problem)であり、これについては様々な学説が提示されている。簡単に示すと、クオリアについて物理的に解明することが可能であるとする心身一元論的な物理主義の立場に反対して、それでは捉えきれない存在としてのクオリアという立場をとった哲学者がネーゲルである。これをより強固に、汎心論的立場から主張を展開したのがチャーマーズである。他方、クオリアが物理的に還元可能であると考えた立場の哲学者の一人がデネットであり、彼はクオリアがなにか新しい、現在の物理学体系から逸脱したものであるという考えを否定し、デカルト以来の二元論性を覆すことを試みている(因みに、駒場の信原幸弘教授もこの立場らしい)。そして、クオリアとその意義について 2011 年に新たな見地を提示したのが、理論心理学者のハンフリーである(後述)。

[補足説明]

●ニーチェ『力への意志』

ニーチェによれば、現在我々のある現象世界と対をなすものは、キリスト教的・プラト

ソ的な「真の世界」ではなく、無限に解釈が可能である「混沌の世界」である。「真の世界」的思考は、惰性(疲労に基づくエコノミー)であり、「混沌の世界」の承認こそが力となると考えた。世界の「無限の解釈可能性」とはすなわちパースペクティブ(遠近法)からなる無限性であり、ニーチェによれば科学もその例に漏れないとされる。

●ディルタイ『解釈学』

・ディルタイの基礎概念

ディルタイは、外(=世界)から感覚的に受容される身振り・表情・言葉などのものを「徴表」と呼び、徴表の意味(内的なもの)を我々が認識する過程を「理解」と呼んだ。また、これと区別して、文献など持続的に残る「生の表示」を技巧的に分析することを「解釈」と呼んだ。

・理解の基本的形式

ディルタイによれば、「理解」の形式には高次なものと低次なもの二種類が存在する。身振り・言葉などの徴表と、その内的なもの(表現されているもの)との規則的關係が共有されている場合、「理解」は基本的な形式で実行される。つまり基本的「理解」は文化コンテクストに依存しており、それを共有するコミュニティの内でのみ、すぐ直接的に「理解」することが可能となる。

・理解の高次形式

文化コンテクストを共有していない場合、他者の心を追体験(想像)する「追構成」によって「理解」が図られる。ここでいう他者とは人間にとどまらず、キャラクターや自然、人工物など多様な存在が対象となりうる。「自己移入 sich in... versetzen」とも呼ばれ、たとえば小説を読むときなどもこれに当てはまるとされる。

●ネーゲル『こうもりであるとはどのようなことか?』

・タイトル『What Is It Like To Be a Bat?』について

「こうもり」であるとは、「こうもりである」という心的状態、すなわちクオリアを伴った心的状態を指す。たとえば、こうもりは超音波を発信してその跳ね返りから距離を測定して、暗闇を飛び回るといった「視覚(ソナー)」を持っているが、それはいったいどのようなことである(it is like to)のか。というように、こうもりの体と脳を所有して諸々の体験をする際に我々は何をどのように感じるのか、結局のところはわからない(とネーゲルは考える)。これが何よりも、客観的な物理事象のみでは捉えられない存在(=クオリア)を示していると考えられるネーゲルの立場を端的に示したのが、このタイトルである。

・経験の主観的性格

ネーゲルはこのような個別の特殊な意識体験を「経験の主観的性格」という言葉で表現している。物理学への「還元主義」の立場ではこの特性は不要であり、それが存在しなくとも還元主義は成り立つ。しかしそのように考えると、「経験の主観的性格」を持つことの

ないロボットなども、人間と全く同じ次元での「心」を持っていることとなる。すなわち、還元主義の立場はこの「経験の主観的性格(≒現象学的特性)」をも物理的に解明しなければ(本当は)いけないのである。

・(他者の)体験の理解と客観的現象学

我々が他者の体験を理解するとはどういうことなのか、ディルタイは表現方法の遠近法(自分と表現方法が似ている他者は基本的理解、遠い人は追構成)から考えたが、ネーゲルは体験主体と自己の(より本質的な)差異から検討を加えている。ある対象が自分と似ている場合は、彼の視点をとって一人称的視点を想起し、「理解」することが可能である。しかし差異が大きければそれだけ、視点の想起は困難となる(ex:こうもり)。

ネーゲルは「主観的性格」について現状では、想像に頼る以外には探求が全くできないとしながらも、これらの記述を部分的にでも可能にする「客観的な現象学」が要求されていると考える。それはたとえばこうもりのソナー体験の記述や、生まれつき全盲の人に色とは何かを説明する記述法の開発などである。この想像に依拠しない説明を形作る客観的現象学は、ネーゲルによれば物理学体系とは別個の、諸体験を記述する体系となりうる。

●デネット『解明される意識』

・ヘテロ現象学

デネットが主に主張することは、脳の神経接続の問題と意識の本質の問題を別次元ととらえる、デカルト的な心身二元論(ネーゲル、チャーマーズなど)の批判である。そもそも、デカルトが提唱した「我」はもともと一人称単数視点であったにも関わらず、それはやがて「我々(一人称複数)」に適用しうるものとして拡張されてきた。デネットはここから一人称視点を問題視する。

三人称視点を重視するとはいったものの、行動主義では「心」がどのような状態であるかについてはまったく仮説を立てられない。デネットは行動主義を避けつつ、客観的物理学と三人称的視点へあくまでもこだわって、そこから心的状態について科学的に認められるような理論として「ヘテロ現象学」を打ち立てた。すなわちネーゲルが言及した客観的現象学を、物理還元主義の立場から構築しようとしたのが、デネットなのである。

ヘテロ現象学においては、体験主体(被験者)の体験から「ヘテロ現象学的世界(被験者の信念で構成された世界観)」を抽出することで、被験者が世界をどのように感じてきたかのあらゆる「意識の流れ」を網羅する。これこそが、中立的描写であり被験者「であるとはどのようなことか(it is like to be)」に対する答えである。

・多元的草稿理論 Multiple Drafts Theory

講義内では言及されていないが、デネットを理解するうえで欠かせない概念なので一応補足しておく。19世紀末にアメリカの心理学者ウィリアムズが「意識の流れ stream of consciousness」を提唱した。これは、人間の意識が静的な配列によるものではなく、動的に、様々な要素が流れるように連なっているという理解であり、『トリストラム・シャン

ディ』などに代表されるように文学表現にも影響を及ぼしている。

デネットの「多元的草稿」はこのような人間の意識理解を、脳を中心とする還元主義に置き換えたものといえる。脳の特定位が意識をつかさどる中央機関として機能しているのではなく、意識は時間的・空間的に多様な drafts の組み合わせによって構成されている。たとえば、一定距離を置いてスクリーンに赤い円と緑の円を並べて、交互に表示させると、あたかも 1 つの円が中点で色変化しながら往復しているように見える。しかし、中点で色が赤から緑に変化した瞬間は、まだ実際には緑の円は映し出されていない。この現象は色をランダムに表示させても同様なので、過去の記憶から脳が馴れたという解釈も成り立たない。つまり、緑の円が表示された後に、緑の円が見えていない過去を「見えていた」と改ざんしている、とデネットは主張する(これを「オーウェルの改竄」という)。これが「多元的草稿」メカニズムを示す一例である。

●哲学的ゾンビ

どこで挿入するか迷ったがここにする。

チャーメーズが物理主義を批判する文脈で用いたことから有名になった概念である。哲学的ゾンビは、我々と外見もしくさも行動もまったく変わらない。私たちと会話し、よく笑うし、怒るし、泣く。脳の神経状態まで人間と全く同様である。しかし哲学的ゾンビには、笑う時の嬉しいという感覚や、怒る時の感覚といったものがない。つまりクオリアを所持していないのだ。哲学的ゾンビは存在しうるのか。

これについては、すでに扱った「中国語の部屋」や、「スーパー科学者マリー」などの思考実験と似たような論点であり、いずれも物理還元主義を否定する立場から引用されるものである。

●ハンフリー『ソウルダスト』

ハンフリーは、意識経験・感覚・クオリア・現象的意識(多くの類語が登場するので、便宜的にクオリアで統一してしまう)が間違いなく存在するものであることが一般の考えであることを前提としつつ、理論家はそこで立ち止まって「クオリアはそれそのものが実態なのではなく、そういう装いだけかもしれない」と考える。ハンフリーによれば、クオリアは巧妙に仕組みられたトリックであり、しかしそうでありながらも、クオリアは進化論的な適合性を持って私たちに備わったものであるという立場をとっている。ハンフリーは著書の中で多様な例を提示しながら論を進めているので、ここでもそれにしたがう。

・盲点

ハンフリーはまず、脳の補正性についていくつか例を挙げて説明している。たとえば、白地の上に一定距離を置いて左右に配列された二つの黒点について、観察者は右目をつむった状態で左目のみで右の点を見る。近づいたり離れたりしながら見ていると、ある一点において左の点が消滅する箇所がある。これが「盲点」である。

盲点は網膜内でありながら網膜のない部分(生理学上必ず存在する)のことであり、恒常的に存在するが、日常生活でそれを感じないのは脳が補正を施しているからである。しかしこの実験では、左の点の焦点が盲点と一致すると、脳は「周りが白いから盲点の箇所も白いのだろう」と解釈して、黒点をないものとして補正して知覚させてしまっていることが、観察者にバレル寸法となっている。

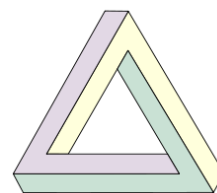
・盲視

受容器官としての眼に異常はないものの、脳の視覚野の一部が損傷して発生した盲目状態を「盲視」という。「見えていないはずなのに見えていると感じる」盲点とは対照的に、盲視については「見えていないと感じるのに実際は見えている」現象が起こりうる。

どうということか。Weiskrantz らは、盲視の患者にいくつかの黒点を提示して「点はいくつ見えますか？」と訊く実験を行った。患者は「見えていないから答えられない」というものの、実験者は当てずっぽうで答えを言わせる。すると、チャンス・レベル以上の正答率であるという結果になった。つまり、盲視の人々は知覚意識、つまりクオリアなしに知覚をしている可能性があるということである。これは、デカルト以来長い間信じられてきた「知覚内容は間違える可能性があるが、知覚しているという意識体験(クオリア)は確実に間違えない」という根本原理を覆しかねない例として、現在注目を集めている。

・不可能図形

右図(Wikipedia より)は、イギリスの科学者ペンローズが考案した「不可能図形」の一種である。このような図形はあり得ないにも関わらず、なぜ今我々はこの図形を知覚できているのだろうか。そもそも「あり得ない」ものは知覚などできないのではないか。



不可能図形はあくまでも二次元図形として存在しているのであり、その限りにおいては存在可能である。しかし我々がその図形を知覚するとき、それは脳によって三次的なものとして処理される。このように脳が補正を行うことによって、「(三次元では)あり得ない」図形が知覚されうるのである。

・グレガンドラム

観察者 A、B がそれぞれ違う角度(パースペクティブ)から、「グレガンドラム」という複雑な形をした立体を観察している。グレガンドラムは A の角度から観察すると「ペンローズの三角形」に見える(グレガンドラムという名をつけたのはおそらくハンフリーだが、ある角度から見るとペンローズの三角形に見える立体は yo***be などを探せば見つかる)。

さて、この時 A の心的表象(A が見ているグレガンドラムの姿。A の知覚の志向的対象)を A_1 、そしてそれに対応する現実の物理世界(実際に存在するグレガンドラム)を A_2 として、同様に B_1 、 B_2 を定義したとする。 A_2 と B_2 は当然一致し、また B_1 と B_2 、そしてグレガンドラムをその他の角度 C、D などから見た場合も $_1$ と $_2$ はやはり一致する。しかし、A の角度だけは、 A_1 と A_2 が一致しないのである(A_1 のような図形は物理界に存在しえないため)。すなわち、志向的対象と現実世界とは一致するとは限らないのである。

・イプサンドラム

ハンフリーはここから、意識一般においてもグレガンドラムと同じような「幻想」が発生しうると展開していく。脳の知覚についてもこのような「幻想」が生じうる以上、志向の対象(クオリア)と現実世界(脳の活動)とは分けて考えていく必要がある。たとえば我々が「赤いもの」を見て「赤い」と知覚するとき、確かに我々の志向的对象(クオリア)としては「赤いもの」がある。しかしこれについても、「感覚的なグレガンドラム」と言ってしまうのではないか。つまり脳の活動とクオリアは、角度 A からグレガンドラムを観察した時のような、幻想的な結びつきしか持たないのではないかとハンフリーは結論付ける。ハンフリーはこのように、「意識・クオリア(X_1)」を脳が作り出した「幻想」であるとしたうえで、クオリアを生む「感覚的なグレガンドラム」としての脳の活動のことを「イプサンドラム(X_2)」と名付けた。

盲点でわかるように、脳というイプサンドラムが補正を加えることで、我々の意識・クオリアは成立している。このことからハンフリーは、クオリアをイプサンドラムが作り出した幻想とする説を補強する。それゆえに、幻想を自己発生させる装置としての「イプサンドラム」をロボットに搭載すれば、理論上は「心を持ったロボット」が成立するかもしれない、というのがハンフリーの立場である。

・進化と意識

クオリアはイプサンドラムが作り出した幻想であると結論付けたうえで、ハンフリーはそれがどのようにして発生したものなのかを、進化論的に考察している。

感覚とはもともと、自らの身体状態を表象するという、生物が生存するのに必須の機能である。つまり最も原初的な感覚とは、外部表現を伴う反応を受信する際のモニターである。つまり感覚とは当初、外部表現に向けて生体を突き動かすための付随的なものであった。

しかし今の人間、いや人間より以前から、生体がそのような感覚刺激についてまったく不随意的に外部表現を見せることはほとんどない。つまり感覚が「外部表現を伴わないもの」として進化(自然淘汰)しているのである。これが更に世代を経て経験を積んでいくことで、感覚は持続的に内面化され、パターン化されていき、これがイプサンドラムと呼ばれるようになった。つまりイプサンドラムとはパターンなのであり、こうして成立したものが「意識」なのである。以上のように、意識とは原初的な感覚が自然淘汰を経て発達していったものだとするのがハンフリーの立場である。

つまり意識・クオリアとは生体を生存させるためにイプサンドラムが生成を続けている幻想である。我々はクオリアそのものを楽しみ、クオリアを通じて世界を楽しみ、そしてそれらを楽しむ自己にも楽しみの気持ちを抱く。クオリアがこのように多様なレベルにおいて喜びをもたらすこと、これ自体が「クオリアが生存にとって有利である」何よりの証拠である。なぜなら、「楽しい」からこそ我々は、クオリアやそれを楽しむ自己を存続させる行為、すなわち「生存」を選択するからである(つまり我々のほうが哲学的ゾンビよ

り生存可能性が高い)。また、ハンフリーに言わせると、幸福一辺倒であるよりも、適度に不幸を取り入れたほうがダーウィン適応度は最大になる、すなわち生存可能性が高い。我々がわざわざ不幸のクオリアを感じ取るのにもハンフリーは意義を見出しているのだ。

因みに石原先生はハンフリーの議論について、クオリアが進化上の結果であったことは認め、機能主義が無視してきたクオリアの議論に取り組んだ点を評価しているものの、そもそも我々の志向的対象(X_1)とイプサンドラム(X_2)の一致不一致は、クオリアそのものの実在の有無とはなんら関係がないと指摘し、それだけでは「クオリア=幻想」説は成り立たないとしている。

【9. 心・脳・社会】

[概略]

最終講では、心や脳のメカニズムとそれがもたらす社会性について解説がされている。

社会性とはすなわち他者理解であるともいえるが、他者理解を考えるうえで最も重要な理論の一つが、「心の理論」である。1978年にブレマックとウッドラフによって体系化されたこの理論は、心的状態を他者に帰属させる能力のことであり、「誤信念課題」などによって検査されることが多い。イギリスの発達心理学者コーエンは、1995年に進化心理学の観点から、心的ステータスを4段階に区切り、その関係から心の理論の整理を行った。一般的に自閉症患者は心の理論を所有していないとされるが、コーエンはここから心の理論のモジュール性を指摘した。

さて、心の理論はどのようなメカニズムに基づく能力なのだろうか。現在は「理論説」と「シミュレーション説」に主に分かれているが、シミュレーション説を支持する研究例が少し前に話題となった。イタリアの学者リゾラッティによる「ミラーニューロン」概念の発見である。ミラーニューロンとは端的に言えば、他者の運動行為の観察を自己の運動行為感覚へと転換させるシステムのことであり、この発見によって、「心の理論」などの他者理解のシステムについて、我々はシミュレーションを行うことによって理解を図っている可能性が補強されることとなった。

最後に授業のまとめとして、哲学者ブロックが取り上げられる。ブロックは意識を「現象的意識(=クオリア)」と「アクセス意識」に分けて考えた学者であり、これは前講で紹介されたハンフリーのような、盲視からクオリアの意義を考察する説に反対するものとして考えられた理論である。石原先生ご自身の考えについては後で説明する。

[補足説明]

●心の理論 Theory of Mind

他者も自分と同じような心的状態を持ちうるものとして考え、そこから他者の行動を予

測したり、説明したりする能力のこと(日本語で考えるならば「心の理論能力」と考えたほうがわかりやすい)である。最初に心の理論を提唱したプリマックによれば、心の理論は、人間は当然所有しているものであり、更に類人猿についても所有の可能性があるとされる(第6講で登場した霊長類研究者のトマセロはこれを否定している)。

・誤信念課題

他者が「誤った」信念を持っていることを理解しないと正答が得られない課題のことで、被験者が「心の理論」を所有しているかどうかを調べる際に用いられる。有名な例として、「サリーとアン課題」がある。『部屋の中にカゴと箱があり、サリーはそのうちカゴの中にボールを入れて部屋を出ました。その後アンがボールをカゴから箱に移しました。サリーが部屋に戻ってきたとき、ボールを取るためにまずカゴと箱のうちどちらを探すでしょうか?』という課題である(調べたほうがわかりやすい)。ここで、自分の心的状態をサリーに帰属させられない人、つまり心の理論のない人は「箱」と答える。

●コーエン『自閉症とマインドブラインドネス』

コーエンは、進化心理学の観点から、生体が他個体を認識する能力を4つに区分した。

・意図の検出(ID)

他個体の意図を把握する能力。ここでいう意図とは「向かってくる(=自分を捕食しようとしている)」「逃げている(=自分による捕食を避けようとしている)」のいずれかであるため、比較的下等な生物でも持っていると言われる。

・視線の検出器(EDD)

他個体がどこを見ているのかを把握する能力。コーエンの実験によれば、「生得的に目の見えない子どもと大人」グループ以外の生体はこの能力を所持している。

・視線の共有(SAM)

ある一個体が注意を向けている対象について、他個体も注意を向けようとする能力。コーエンによれば、下等動物・9か月以下の人間・一部の自閉症患者以外はこの能力を所持している。

・心の理論(ToMM)

すでに説明したとおりである。コーエンの実験によると、発達障害がある人でも精神年齢が4歳以上であればこの能力は所持されているが、自閉症患者については精神年齢を問わずこの能力に難のある人が多い。コーエンはここから、心の理論がモジュール性をもった「特化能力」であること、つまり自閉症は症状のメカニズムとしては盲人と同じであることを主張し、彼らの症状を「mindblindness」と呼んだ。

●モジュール性 Modularity

アメリカの哲学者フォードは、脳・神経のメカニズムについて、コンピュータの情報処理過程と同様に「入力系」と「中央過程」の二種に分けることができると主張した。

・中央過程

全体的であり、特化していない認知過程のこと。フォーダーによれば「入力系以外の認知過程」のことで、何らかの形で信念体系全体に関与している過程を指す。

・入力系

領域特異性・情報遮断性を持った感覚系のこと。たとえば視覚系については、それぞれの器官は視覚のみに関与する(他の感覚系から)独立した特化器官であり、さらに聴覚系とは互いに情報を遮断している。このように、それぞれが特異性と遮断性をもった感覚器官が集まってひとつの感覚系を成している構図について、フォーダーはこれらの器官を「機能単子(モジュール)」と呼んだ。

つまりモジュール性をもった能力とは、特化された、ある特定の部位でその能力を司っているような部位が存在する能力のことを指す。たとえば視覚のモジュールに何らかの異常があれば、その個体は「盲 blindness」となる。コーエンは、自閉症症例の研究から、心の理論についてもモジュール性が認められると主張しているために、相手の心が読めない障害という意味で「心盲 mindblindness」という呼び方を用いている。

●心の理論のメカニズム

・理論説 Theory-Theory

他者の心は、理論的・理性的な推論によって構築されていくという考え方。そもそも心の理論を最初に提唱したプリマックは、この能力が理論性を基盤としていていると考えたために、心の「理論」という呼称をつけたとされる。

理論説が前提としているのは、そもそも我々が他者の心に直接アクセスすることが出来ないという点である。そうである以上、我々は一般的な法則に基づいて推論を試みるしかない、とするのが理論説の立場である。

・シミュレーション説 Simulation-Theory

他者理解について、理論的な推察ではなく、自分の心の一部を作動させ、それをモデルとすることで他者の心をシミュレートしているのだとする考え方。理論説の前提である「相手の立場に立つことなど出来ない」に反対し、相手の立場から物事を感じるシミュレート能力が我々に備わっていることを主張している。

●ミラーニューロン

リゾラッティは 1996 年、サルの脳に電極を埋め込み、サルが自分でエサを取る際の脳の状態と、実験者がエサを拾い上げる様子をサルが見ている際の脳の状態とを比較した。すると、いずれの時においても同じ神経細胞が活動していることを発見した。この「鏡」のように自己の活動と他者の活動観察の両方で働く神経細胞は「ミラーニューロン」と名付けられた。

ヒトの脳について、サルと同様に細胞単位の観測を行うことは困難であるが、ヒトにつ

いても複数の要素が絡み合っサルよりも複雑な機構としての「ミラーニューロン・システム」が存在するであろうことが明らかになっている(らしい)。このシステムは、他者の行為を記述(表象)する情報と、自己の運動行為の情報が、同じニューロンで統括されている可能性を示しており、すなわち観察しただけでそれが自己の運動行為に転換されうること、つまり「見ただけで経験した気になる」システムであることがいえる。

ミラーニューロンが発見されたことによって、我々は理性的推論の他にもシミュレーションによって他者理解を遂行しているという、2つの説の両立が考えられることとなったが、ミラーニューロンについてはいまだ研究途上にあるため、なんともいえない(らしい)。いずれにせよ、人間が生きるうえで欠かせない「経験の共有」という社会性が育まれる素養が、神経科学的に説明されうる可能性は、特筆に値することであるといえる。

なお現在では、サイコパシー(精神病質者)の診断的特徴である「共感障害」について、ミラーニューロン・システムに関する器質的な異常からこれを説明しようという動きも存在する。Fecteau らによる実験では、人間の手に針が刺さる様子を記録したビデオを被験者に見せた時の反応電位が、「冷酷さ尺度(他者に共感しない指数)」と負の相関にあることが実証されている。

●社会脳

「心の理論」を含む人間や霊長類の社会性・社会的行動について、ニューロンレベルの活動との関係から研究することを提案したイギリスの進化人類学者ブラザースによって、1990年に初めて使用された言葉である。1998年には同じイギリスのダンバーによって、霊長類の脳の相対的大きさを、その複雑な社会システムへの適応の必要性による淘汰から説明しようとする「社会脳仮説」が提唱された。

このような流れに伴って、社会神経科学と呼ばれる学問領域が90年代から勃興するが、これは必ずしも「社会的・心理学的プロセスを脳の特定領域における活動にマッピングする」ことを指すわけではない。ここではむしろ、人文社会科学と自然科学が双方の理論に互いに情報を与えるような、横断的な取り組みの架け橋となる可能性が期待されている。

●ブロック『Behavioral and Brain Sciences』

前講で取り上げられた盲視の例から、現象的意識(クオリア)を欠くが一定の知覚が可能であるケースが明らかになった。しかし、実際には盲視の患者は「何も見えていない」と感じているため、知覚に基づく行動が出来ない。ここから一部の研究者は、クオリアは行動の制御のために存在するのではないかという説を提唱している。

ブロックはこの説に反対し、現象的意識と区別されるものとして「アクセス意識」の存在を主張している。ブロックによるとアクセス意識とは、思考・推論や発話行為の理性的なコントロールを統括するものであり、つまり「現象的意識の報告を基礎づける認知的機構」を指す。ブロックは、このように「現象的意識の報告」は現象的意識そのものとは区

別されるべきであると主張し、現象的意識の神経基盤そのものはまた別にあるものと考えている。

第1講でブレイスウェイトが取り上げていた「アクセス意識」の概念とは異なる面もあるが、ブロックはランドマンの認知心理学実験を援用しながら「アクセス意識の容量以上のものを現象的意識は持っている」と主張している。つまり、現在の神経科学で解明可能なアクセス意識とは別の存在として現象的意識をとらえている。

●石原先生のまとめ

講義内容を見た限りでは、先生はどちらかということと二元論的立場を採用しているようである(が、あまり明確にその立場を主張しないのは、職場の先輩である信原先生が一元論派だからなのか)。講義の最後でも「意識とは現象的意識を伴うものである」というまとめ方をしており、さらに「心」は、神経科学によって完全に置き換えられることはなく、現在理解されているような形で「実在」すると考えてよいだろうとしている。

それ以外の論点については先生としては明確な立場を主張していない(ロボットについて、植物状態について)。脳神経倫理(Neuroethics)という言葉を引き合いに出して、主体の「意識」と責任の問題について再度言及しているので、その辺りが試験に出るのかもしれない。よくわからない。

・・・以上でまとめは終了です。

前半と後半はやっつけ仕事感がにじみ出ていますね。

低クオリアですみません。試験がんばりましょう。

なにかあったらメールなり何なりしてください。

2012/07/13 (Fri)

文三 17 組