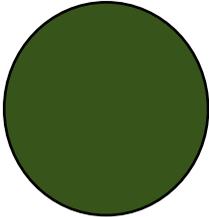


このシラバスには、受講開始から単位を取得するまでに必要なすべての情報が記載されています。再配布はしませんから大切に保管して、つねにチェックするようにして下さい。



## 情報哲学の展開2 -心とロボットの現代哲学入門- teaching plan

戸田山和久 情報科学研究科棟 210 号室 [todayama@info.human.nagoya-u.ac.jp](mailto:todayama@info.human.nagoya-u.ac.jp)  
月曜日 1 限

注意 この授業は、人類がまだよく分かっていないことについてのものです。この授業で扱う問題に関しては、まだよく分かっていないんだけど、何がよく分からないのかということはだんだん分かってきて、それが分かるためには何をしなければならぬかということがぼんやり分かってきた、という状況に我々はいまいます。いろんな考えが試されて、比較されて、勝ったり負けたり、ダメになったり、かと思うと意外なところで復活したりします。というわけで、あれこれ考えるのが好きな人向きです。ごちゃごちゃ面倒くさいことはいいから、早く答えを教えてちょ、という人には絶対に向かない授業だということをあらかじめお断りしておきます。

この授業で考えていく問い

(1) この世界が物質（物理的存在）だけからなる世界であって、それ以外の実体が存在しないのだとすると、心というどう考えても物質ではない存在は、どのようにしてこの物質的世界に座を占めることができるのだろうか。それとも、心があるというのは幻想なのか？

(2) この問いを逆に考えてみよう。心がどうにかして物質によって実現されているのなら、少なくとも原理的にはわれわれは人工的に心を作り出すことができるはずだ。しかし、どうすれば心、あるいは心を持つロボットを作ることができるのだろうか。それどころか、どんなロボットができれば、そいつは心を持っていると言えるのだろうか。つまり、われわれはどんなものをつくれればよいのだろう。

### Schedule

10月3日 ガイダンス

10月17日 問題を立てる ~What is "Mind-body problem"?

Part I で扱う問題は「心身問題 (mind-body problem)」と言われる伝統的問題の「現代版」だ。そこでまず、心身問題を哲学の問題としてきちんと立ててみよう。そのためには、心身問題がどのような枠組のもとで「問題」として現れてきたのかを歴史的に押えておくことが有効だろう。登場人物は 17 世紀の哲学者デカルトだ。というわけで今回だけは「現代」哲学の話ではないけど勘弁。心身問題と、外界の存在についての懐疑論と「われ思うゆえにわれあり」という超有名なフレーズのとの関係をすっきりさせる。その上で、心身問題が現代においてもまだ解かれていない超難問であることを納得してもらおう。なんせ、犬にはどんなにがんばっても量子力学の問題が解けないのと同じように、人間の知的能力では心身問題を解くことは不可能なのかもしれない、という人もいるくらいだから。



Part I 物質だけの世界に心を描き込む試み

10月24日 機械の中の幽霊～哲学的行動主義とその限界

心身問題を「問い」たらしめている前提条件のひとつは、心が物質とは異なる何らかの実体であると考えることが謎に思われてくる。これが心身問題だった。物質とは別の実体として心を想定することはもちろん現代科学の立場とは相容れない。そこで、現代の心身問題は次のようなものになる。この世界に存在するものはつまるところ物質だけであり、この世界で生じるさまざまな現象は究極的には物質の振る舞いについての科学で説明されるのだとするなら、なぜ心について語る表現に意味があり、心について研究する科学が成り立っているのか？ 現代版心身問題とは、要するに物質だけが存在するという世界観と、われわれは心があると思っていて、心についての記述が有意味で心の科学が存在すると思っっているという事実とをどのように調和させるかという問題だ。この現代版心身問題に対する最も初期の解答の1つが論理的行動主義である。しかし、この考え方はうまくいかなかった。それはなぜなのかを考える。

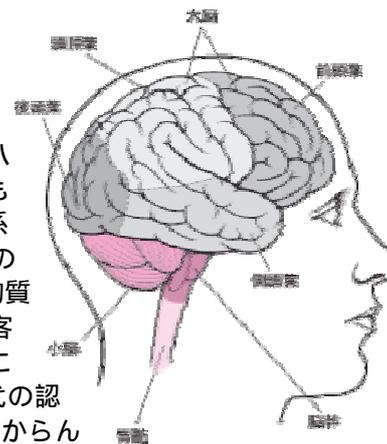
10月31日 心と脳を同一視できるか～心脳同一説（還元主義）の意外な難点

心はどこにあるかと問われれば、ほとんどの人が脳にあると言うだろう。そしてこのことには良い証拠がある。しかし、心は脳にあるとか、心の働きは脳によって担われる、と言うことはすなわち心は脳に他ならないということではない。二元論者のデカルトですら、「脳は精神の座である」ことを認めていたのだから。ここでは、心は脳に他ならないとする哲学的見解を検討する。心脳同一説（mind-body identity theory）は、さらに一歩進んで、心の状態というものは脳の状態と同一だという具合に、完全に心を脳と同一視して、心の状態を脳の物理的状态に還元する。しかし、ある意味で常識的であり現代科学ともじっくりいきそうなこの立場は、つきつめるとかなりの困難を含んでいる。それを明らかにしていこう。

11月7日 休講

11月14日 プログラムとしての心～多重実現可能性のテーゼと機能主義（非還元主義的唯物論）

脳をコンピュータになぞらえてみよう。そうすると、心はそのハードウェア上で走るソフトウェア、つまりプログラムに相当するものだと考えられる。機能主義は、我々の心の状態を脳という物質系において実現されている抽象的な機能的状態（つまりプログラムの計算状態）として捉える。機能主義は、この世界に存在するのは物質だけという世界観の中で、心についてのさまざまな語りを最大限客観的で文字通り正しいものとして位置づけようとする、現在のところ最も洗練された立場であり、心理学と人工知能を中心とする現代の認知科学の基盤をなす哲学だと言える。...と言っても何のことやら分からんでしょ。そこでまず機能主義とは何かについて解説しよう。



11月21日 心のモジュール性～機能主義の展望と限界・そして還元主義の逆襲

今回の講義は2つのパートに分かれている。前半では、機能主義のいろいろなバリエーションを紹介し、機能主義が認知心理学と人工知能研究の基盤を与える哲学的立場だというのはどういう意味かを明らかにする。しかし、これほど有望そうな機能主義にも大きな困難がある。後半では、フレーム問題をはじめとする機能主義に対するいくつかの重要な批判を扱う。とりわけ、心的因果ということをもじめに考えようとする、とっくに退けたはずの還元主義が息を吹き返して逆襲してくる。【Take-home exam 1 配布】

Part II 物質だけの世界から心を追い出す試み

11月28日 人類は「あ～僕のC-fiberが発火している」と言うようになる？

～消去的唯物論 vs. folk psychology

Part I では、物質だけの世界の中に心に対応するする何か（行動、脳の物理的状态、脳で実現される計算状態）を探し求める試みを紹介した。しかし、このように物質だけの世界に心を描き込む試みはなかなかうまくいかなかった。となると、次のような考え方が出てくるだろう。我々が心について語っていることは

間違っているかナンセンス、せいぜい不正確なので、世界についてのまじめな記述や説明からは排除すべきなのではないか。つまり、心について語る言語は捨てるべきであり、それらは単に時代遅れの「民間理論 folk theory」の名残にすぎず、「悪霊」や「かみなりさま」が世界についての説明から消え去ったように、科学の進歩に従って「痛み」や「信念」、「怒り」といった語彙も我々の日常言語から消えてなくなるのではないか。このおそろしくラディカルな主張が消去的唯物論と言われる。まず、消去的唯物論の主張に耳を傾けよう。

12月5日 消去的唯物論の主張をマジに受け止めなければならないわけ

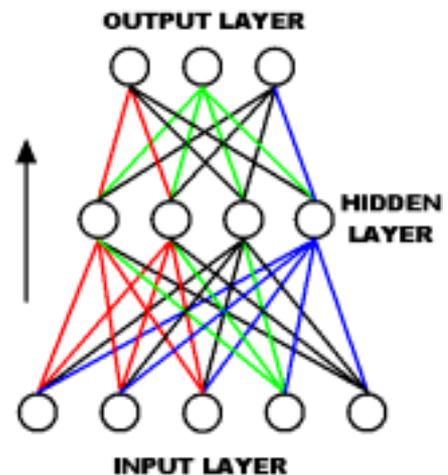
～コネクショニズム vs. 思考の言語仮説

消去的唯物論が、これまでの科学の歴史からの単なる未来予測にすぎないのだったら、その主張がいかにラディカルであっても、われわれはそんなにそれを気にする必要はない。しかし、実態はそうではない。消去的唯物論には、認知科学・脳科学・コンピュータサイエンスからの強力な援護射撃がある。つまり、コネクショニズム、またの名をニューラルネットワーク・モデルという考え方だ。これは、人間の心的な情報処理を何らかの言語のような構造を持った対象（思考の言語）に対する規則に基づいた変換操作として捉える、認知科学の主流派に対するアンチテーゼになっている。そこで、コネクショニズムと消去的唯物論との関係をきちんと理解することを目指そう。

【Take-home exam 1 提出締め切り】

12月12日 道具としての心～デネットの道具主義的心の哲学

消去的唯物論の議論は、その前提を認めるかぎり、結論を否定するのは非常にむずかしいように思われる。しかしながら Daniel Dennett は、消去主義者と前提を共有しつつも結論を異にする。つまり、かれは、心的状態に対応する脳状態の存在を認めず、とりわけ信念や欲求が何らかの形で脳の中にリアルに実現されていて、体を動かすというような因果的力能をもつということを認めない。したがってデネットは、信念や欲求という単位で心を語る民間心理学は世界の描写としては間違っていると考える。しかし一方で彼は、民間心理学が将来の神経科学によって取って代われ消滅するであろうということは認めない。このような立場がどのようにして維持可能かを見ていこう。【Take-home exam 2 配布】



Part III 心を科学的世界像と調和させることがなぜこんなにも難しいのか：志向性と意識



12月19日 心の内容は何が決めているのか～志向性を自然化する

Part II までの議論を通じて、心を科学的世界像と調和させることは意外に難しいことがわかったらう。いったいなぜそんなに難しいのかを反省してみよう。まず第一に、心は志向性を持っているからだ。これは心の状態は内容を持っている、と言いかえてもよい。たとえば、いまキミは「お昼にモヤシそばを食べたいな」と思っているとす。この欲求はモヤシそば「について」のものだ。こうした欲求がキミの脳の何らかの状態で実現されているということは認めよう。しかし、だとしたら脳という物質系の物理的状态にすぎないものが、どうして、モヤシそば「について」のもので

ありうるのだ？モヤシそばは、明らかに脳の中には存在していない。このように、心の状態は内容を持ち、内容を持つから、心の外部の何か「について」のものとなりうる。こうした「について」性、つまり志向性をもつ状態は、他の物理状態にはちょっと見あたらない。というわけで、心の状態が結局のところ物理状態

なのだとしても、それはものすごくユニークな状態だということになってしまう。そこで、心の状態という物理状態がどうして内容・志向性をもてるのかを、中国語の部屋、双子地球、スワンプマンといった思考実験を通じて明らかにしよう。

1月16日 心についての最大の謎～意識とクオリアの問題

心を科学的世界像と調和させることがなぜ難しいのか。第二の理由は、心、とりわけ意識は質をもっているからだ。たとえば、いまキミはお昼に、名大南門を出たところにある中華料理屋「大雅」で、好物のモヤシそばを食べている。上にかかった具と麺を箸でつまみ上げてまず一口。麺のしこしこした歯触りと絶妙な火の通り具合のモヤシのしゃきしゃき感、それがあんかけのとろ～りとした舌触りと相まって最高。次に、レンゲで澄んだスープを掬って一口。こんどは鶏ガラのダシのコクに、醤油の味と胡椒の香りが口中に広がる…。こうしたモヤシを食べたときのしゃきしゃきした感じ、鶏のダシの、ビーフコンソメとも鰹だしとも違うあの独特の味、こうした感覚の持っている質の側面（これをクオリアと言う）は、モヤシそばを食べているキミの脳で生じている物理化学的变化をどんなに詳しく記述しても、そこから抜け落ちてしまうように思われる。だとするなら、物理化学的な世界の記述はクオリアを扱えない点で本質的に不完全だということになってしまうのだろうか？クオリアの存在に訴えて唯物論的な心の哲学を批判する2つの重要な議論、(1) 逆転クオリアの議論、(2) 知識論法、を順に検討する。【Take-home exam 2 提出締め切り】



1月23日 ゾンビでどこがいかなのか？～賢いロボットと表象なき認知

われわれの心はたしかにクオリアに満ちた表象を持っている。しかし、そもそも認知には、心には、表象が不可欠なのだろうか。そうではないという考え方がある。MITのロボット工学者であるRodney Brooksは、知的なロボットをつくるためには、外界の表象を持たせることはじつはあまり効果的ではないということを中心主張して、「表象なき知性」というスローガンを提唱した。心は、外界の様子を表象して、それに基づいて行動をコントロールするというわれわれの常識に挑戦するこのスローガンは、逆に「表象」とは何かという反省を促すことになった。「表象なき知性」とそれをさらに展開した認知の力学系理論の意義を探る。

1月30日 そもそも心は頭の中にあるのか？～環境主義の心の見方

この授業の締めくくりに、心は頭（脳）の中にあると考えるべきなのか、という問いを考えてみよう。2365×564という計算を筆算で行っている人を考えてみよう。その計算は、心が行っているのだろうけど、その人の脳の中に2365という数の表象があると考える必要はないのではないか。2365は、紙の上にある。そうすると、この計算というプロセスは、脳と手と紙と鉛筆からなるシステムが行っていると考えた方がよい。心は脳だけでなく身体や外部環境にもまたがって存在しているという考え方を検討する。【Take-home exam 3 配布】



2月20日 【Take-home exam 3 提出締め切り】

COURSE REQUIREMENTS

各パートの最後にそれぞれ Take-home exam を行う（計3回）。Take-home exam は、授業中に配られた問題について、自宅で解答を作成する形式のテスト。問題は11月21日、12月12日、1月30日に配布する。Take-home exam は各35点満点で計105点。全部で60点以上の得点を獲得することが単位取得の条件とす

る。これ以外の期末試験などは行わない。60%～69% C、70%～84% B、85%～ A  
Take-home exam は、講義に毎回出席してその都度疑問点を解消していればとても簡単だが、そうでない場合にはものすごく苦労することになる。どちらの道を選ぶかは自分の責任。

#### 提出方法

- 【Take-home exam 1 & 2】指定された日までに教室で提出。その際、次の点に注意。
- ・A4 のレポート用紙に書いてホチキスで綴じる。
  - ・その際、必ず1部コピーをとって手元に保存しておくこと（万が一、提出したレポートが紛失したときのための用心）。
- 【Take-home exam 3】教室で受け取ることができないので、指定された日までに電子メールで提出。その際、次の点に注意。
- ・アドレス：[todayama@info.human.nagoya-u.ac.jp](mailto:todayama@info.human.nagoya-u.ac.jp) まで提出すること。
  - ・添付ファイルではなく、課題はメール本文に書き込んで送信すること。
  - ・メールの「Subject」の欄に忘れずに「mind+学生番号+氏名」と記入すること。
- 例) mind050500529 滝野智
- ・受信を確認した旨の返事を出すから、提出後1週間をすぎても返信がない場合は戸田山にコンタクトをとること。

---

#### その他の注意

授業は補助資料としてハンドアウト（プリント）を使って行う。ハンドアウトは授業のときにしか配らない。ということは、授業をさぼったからといって研究室にもらいに来てもらえないよ、ということ。（それにいちいち対処していると仕事にならんよ。）ただし、紙資源有効利用の観点から、すべての講義のハンドアウトの余りを研究室の前においた段ボール箱のなかにいれておくから、その中を探せば運良く見つかるかもしれない。そこになければあきらめて下さい。

とは言うものの、ハンドアウトだけ読んでも授業の内容はおそらくわかりません。出席はとりませんが、出席して話を聞いていなければ Take-home exam に対応することはできないと思う。欠席は自己責任で。

スケジュールを見ると、一回読み切りの授業に思えるかもしれませんが、それは甘い。それまでの話を聞いていないと、ほとんど何を言っているのかわからないと思います。

授業中、「わからんぞ」と思ったときにはいつでも中断して質問して下さい。また、授業終了後も少し教室に残っていますから、授業中に質問しそびれた場合は質問歓迎。

私語、遅刻と途中退室は他の人の学習権を侵害しますからやめてください。

---

#### 読書案内

それぞれのトピックに関する読書案内は、毎回ハンドアウトの最後につけておきます。ここでは、心の哲学の全般に関する入門書を挙げておきます。

柴田正良『ロボットの心：7つの哲学物語』講談社現代新書

著者の柴田氏は名古屋大学のOBで、「大雅」のもやしソバのファン。ロボットに心を持たせることは可能という結論を目指して、脳科学やロボット工学、哲学を総動員して議論する。それぞれの章の冒頭に、SFもどきのショートストーリーが配置されていて、楽しく読める。まずは、この本からトライしてみたらよいでしょう。

信原幸弘『シリーズ心の哲学：人間編』勁草書房

信原幸弘『シリーズ心の哲学：ロボット編』勁草書房

柴田本をクリアしたら、もうちょっと本格的なこの2冊に進もう。心の現代哲学についての主要問題が、もれなく扱われている。寝転がって読めるような本ではないが、この2冊を読了したら、心の哲学のだいたいの論点がマスターできる。