

wstęp do filozofii umysłu i sztucznej inteligencji

5.10.2023

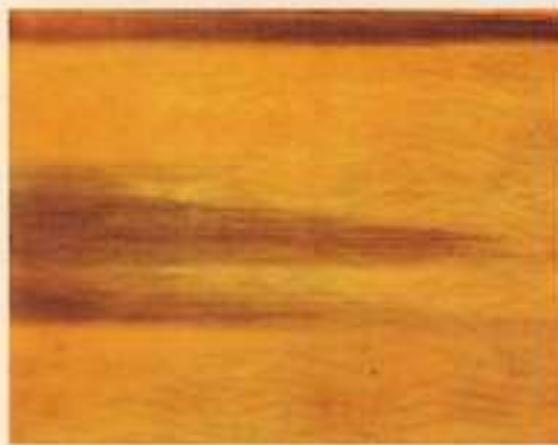
WYKŁAD 1



Paul M. Churchland

Mechanizm rozumu,
siedlisko duszy.

Filozoficzna podróż w głąb mózgu



filozofia umysłu a dyscypliny pokrewne

- ▶ psychologia
- ▶ neuronauki
- ▶ kognitywistyka
- ▶ sztuczna inteligencja

teoria umysłu

- ▶ wyjaśnianie działania poprzez przypisanie odpowiednich stanów mentalnych

Eryk *nudzi się w naszym towarzystwie, więc chciałby jak najprędzej je opuścić, gdyż stęsknił się za Henrykiem i myśli tylko o nim.*

(teoria umysłu = mentalizacja, nastawienie intencjonalne, system czytania w umyśle)

teoria umysłu



teoria umysłu



teoria umysłu

Shakey
(1966-1972)



Charles Rosen
(1917-2002)

problem umysł–ciało



problem umysł–ciało

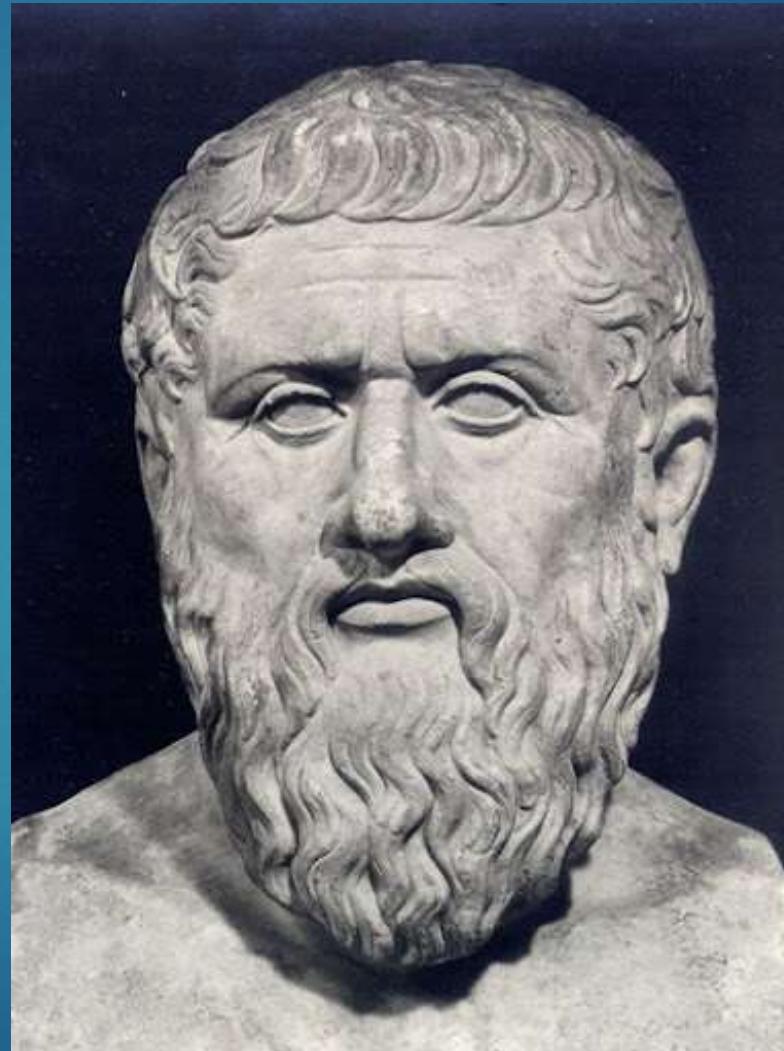
- ▶ jaką jest relacja między zjawiskami **umysłowymi** a **fizycznymi**?
 - ▶ subiektywne – obiektywne
 - ▶ prywatne – publiczne
 - ▶ jakościowe – ilościowe
 - ▶ celowe – mechanistyczne
 - ▶ nieprzestrzenne – przestrzenne
 - ▶ intencjonalne – nieintencjonalne
- ▶ problem przyczynowości mentalnej

filozofia starożytna

- ▶ filozofia przedplatońska – materializm



Platon
(ok. 429-347 p.n.e.)



filozofia starożytnej

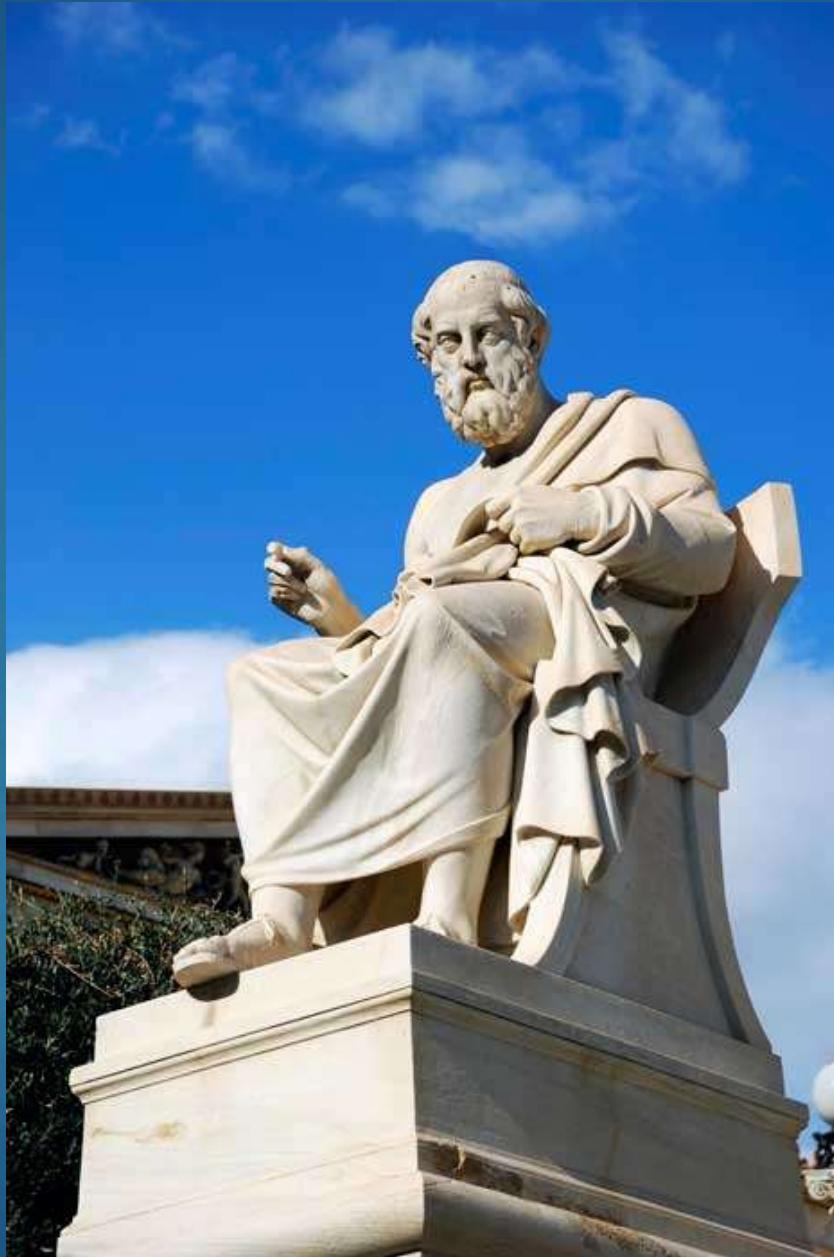
- ▶ filozofia przedplatońska – materializm
- ▶ Platon – teoria duszy
 - ▶ metempsychoza (preegzystencja duszy) – anamneza
 - ▶ psychologiczna koncepcja duszy

filozofia starożytna

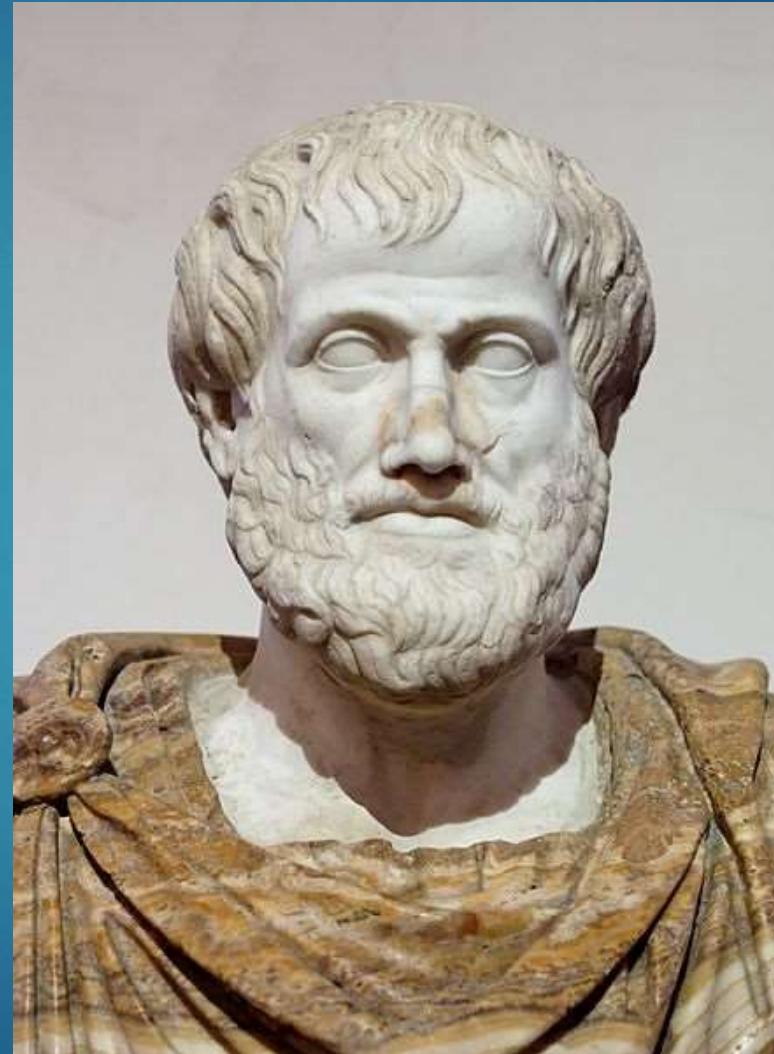


filozofia starożytna

- ▶ filozofia przedplatońska – materializm
- ▶ Platon – teoria duszy
 - ▶ metempsychoza (preegzystencja duszy) – anamneza
 - ▶ psychologiczna koncepcja duszy – trzy części: pożądliwa, popędliwa, rozumna



Arystoteles
(384-322 p.n.e.)



filozofia starożytna

- ▶ filozofia przedplatońska – materializm
- ▶ Platon – teoria duszy
 - ▶ metempsychoza (preegzystencja duszy) – anamneza
 - ▶ psychologiczna koncepcja duszy – trzy części: pożądliwa, popędliwa, rozumna
- ▶ Arystoteles – traktat *O duszy*
 - ▶ hylemorfizm – złożenie materii i formy
 - ▶ dusza jako forma (nieożywionego) ciała
 - ▶ trzy części duszy: wegetatywna, zmysłowa, rozumna

filozofia nowożytna

- ▶ oddzielenie filozofii od teologii
- ▶ odejście od scholastycznego rygoryzmu
- ▶ filozofia uprawiana również w językach narodowych
- ▶ większa koncentracja na pytaniach epistemologicznych
- ▶ racjonalizm vs. empiryzm
 - Kartezjusz, Spinoza, Leibniz, Malebranche, Pascal...
 - Locke, Hume, Berkeley, Bacon, Hobbes...

Kartezjusz (René Descartes) (1596–1650)



filozofia umysłu współcześnie

- ▶ termin „filozofia umysłu” (*philosophy of mind*) popularny w filozofii anglosaskiej od lat pięćdziesiątych
- ▶ wielka debata inspirowana częściowo rozwojem informatyki (*computer science*)
- ▶ najważniejsze problemy: umysł–ciało (przyczynowość mentalna), świadomość, intencjonalność, wolna wola, silna sztuczna inteligencja (*artificial general intelligence*, AGI)



wstęp do filozofii umysłu i sztucznej inteligencji

12.10.2023

WYKŁAD 2

Kartezjusz (René Descartes) (1596–1650)



problem umysł–ciało

- ▶ jaką jest relacja między zjawiskami umysłowymi a fizykalnymi?
 - ▶ subiektywne – obiektywne
 - ▶ prywatne – publiczne
 - ▶ jakościowe – ilościowe
 - ▶ celowe – mechanistyczne
 - ▶ nieprzestrzenne – przestrzenne
 - ▶ intencjonalne – nieintencjonalne

dualizm substancji



dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała

„I chociaż mogę posiadam ciało, które ze mną jest bardzo ściśle złączone, niemniej jednak jest rzeczą pewną, że zaiste ja jestem czymś różnym od mego ciała i bez niego mogę istnieć – ponieważ z jednej strony posiadam jasną i wyraźną ideę siebie samego jako rzeczy myślącej tylko, a nie rozciągłej, a z drugiej strony wyraźną ideę ciała jako rzeczy rozciągłej tylko, a nie myślącej”.

Medytacje o filozofii pierwszej

dualizm substancji

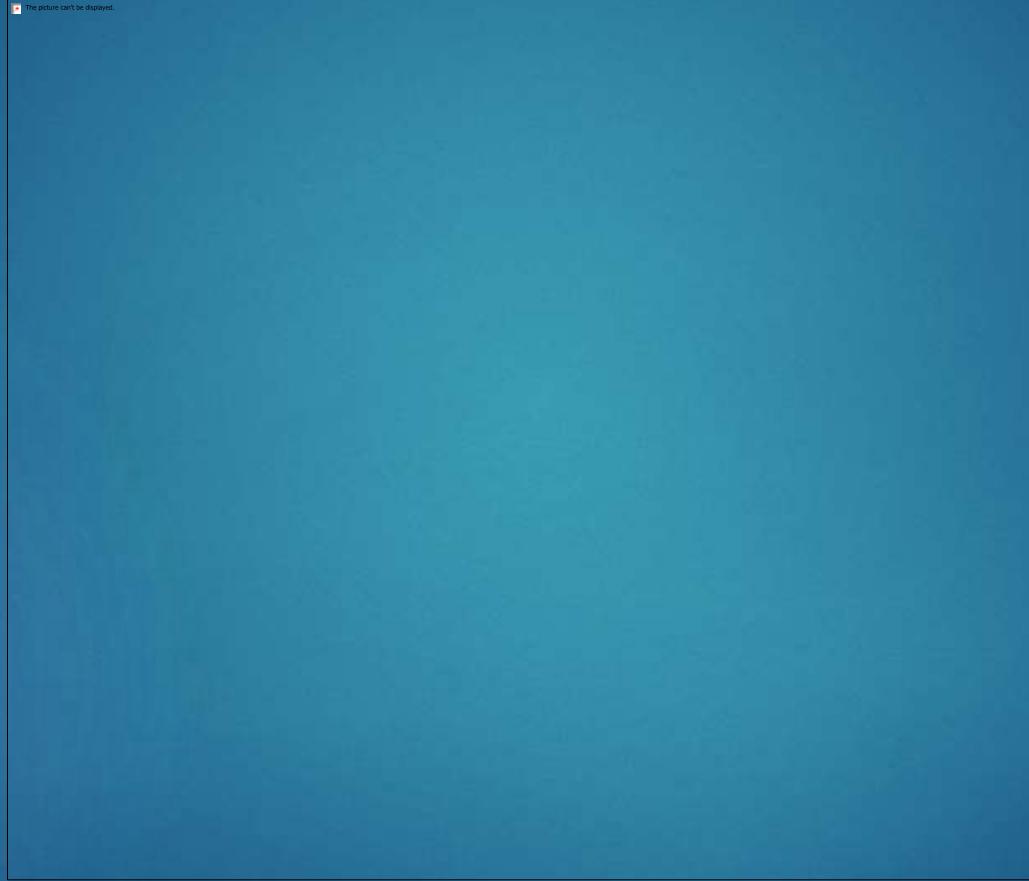
- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
 - ▶ *cogito ergo sum*

dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
- ▶ mechanistyczne rozumienie ciała
 - ▶ fizykalizm
 - ▶ związki przyczynowe

dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
- ▶ mechanistyczne rozumienie ciała
- ▶ dusza = niepodzielna, świadomy umysł dostępny tylko od wewnętrz
- ▶ interakcja duszy i ciała



*Figura
XXXV.*

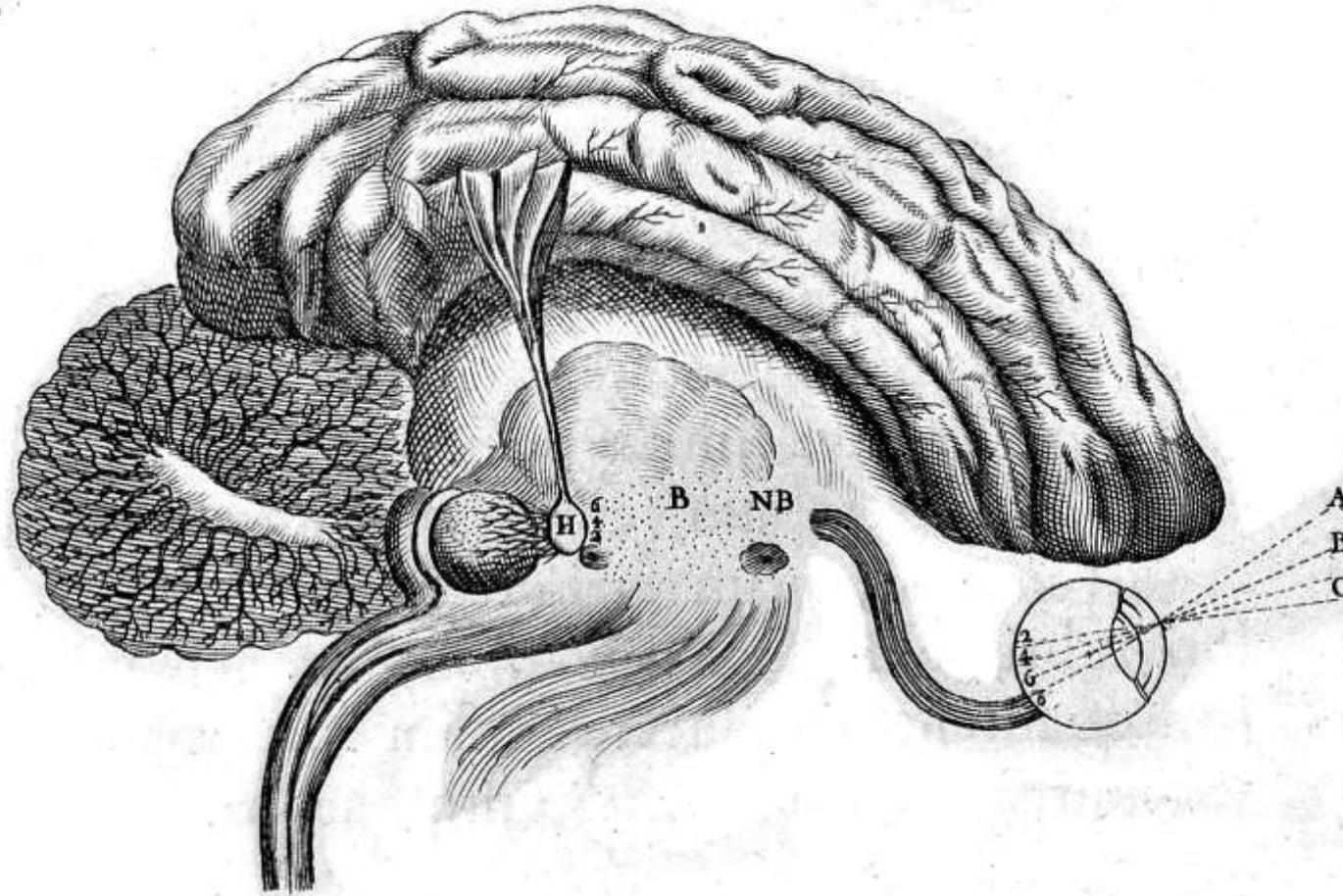
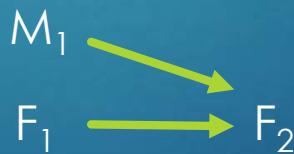


Fig. XXXV.

dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
- ▶ mechanistyczne rozumienie ciała
- ▶ dusza = niepodzielna, świadomy umysł dostępny tylko od wewnętrz
- ▶ interakcja duszy i ciała: M – zdarzenie mentalne, F – zdarzenie fizyczne
 - ▶ interakcjonizm



dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
- ▶ mechanistyczne rozumienie ciała
- ▶ dusza = niepodzielna, świadomy umysł dostępny tylko od wewnętrz
- ▶ interakcja duszy i ciała: M – zdarzenie mentalne, F – zdarzenie fizyczne
 - ▶ interakcjonizm
 - ▶ paralelizm

$$F_1 \longrightarrow F_2$$

$$M_1 \longrightarrow M_2$$

dualizm substancji

- ▶ możliwe jest pojęcie ciała bez duszy i duszy bez ciała
- ▶ metodyczne wątpienie
- ▶ mechanistyczne rozumienie ciała
- ▶ dusza = niepodzielna, świadomy umysł dostępny tylko od wewnętrz
- ▶ interakcja duszy i ciała: M – zdarzenie mentalne, F – zdarzenie fizyczne
 - ▶ interakcjonizm
 - ▶ paralelizm
 - ▶ epifenomenalizm



dualizm substancji

„(...) dualizm substancji traktuje się – niczym szpicrutę, tabakierę i patefon – jako ciekawostkę wyłącznie historyczną, a jego obrońców przyrównuje się do wspomnianych przez Kartezjusza szaleńców, którym się zdaje, że »mają głowę z gliny, albo że całe ich ciało jest dzbanem, lub że są wydęci ze szkła«. Dualiści to tacy szaleńcy, którym się wydaje, że są wydęci z ektoplazmy”.

Marcin Iwanicki, *Dualizm psychofizyczny*, [w:] *Przewodnik po filozofii umysłu*, red. M. Miłkowski, R. Poczobut

dualizm substancji



dualizm substancji

- ▶ Ja vs. moje ciało
 - ▶ mój mózg jak mój żołądek
 - ▶ „Po konsultacji z farmaceutką, a przed zapoznaniem z treścią ulotki, wybrałem preparat najtańszy, **żeby mój mózg nie pomyślał**, że wpadam w panikę” („Dziennik Polski” 2008)
 - ▶ „Papa Stamm mawiał: »Uważaj, żeby **twój mózg** nie stał pod wpływem emocji«” („Dziennik Zachodni” 2008)
 - ▶ „Dlatego jeśli chcesz, aby **twój mózg pracował** pełną parą, **sięgaj częściej po łososia, sardynki, śledzia”** („Dziennik Łódzki” 2003)

dualizm substancji

Does Belief in Dualism Protect against Maladaptive Psycho-Social Responses to Deep Brain Stimulation? An Empirical Exploration

Jason Shepard & Joshua May

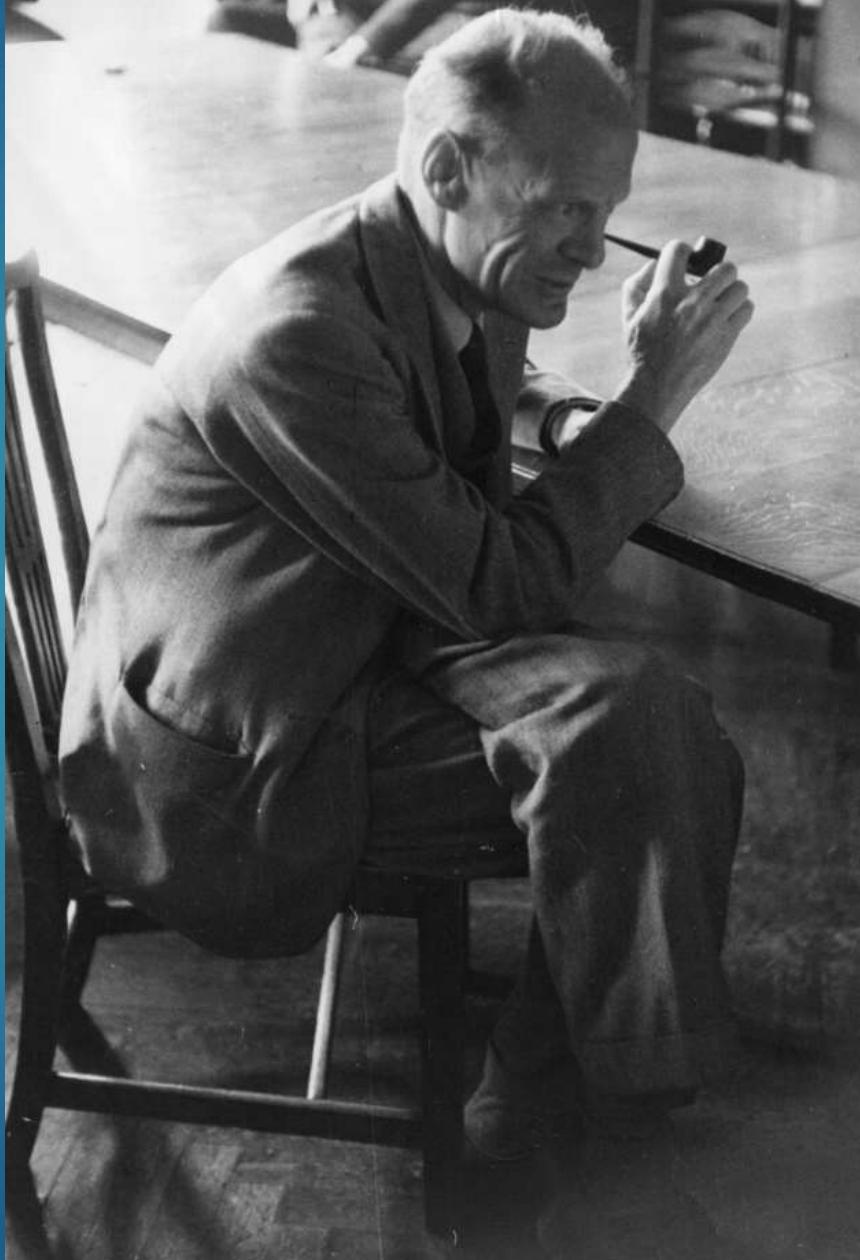
Published in the *American Journal of Bioethics: Neuroscience* 5(4): 40–42 (2014).

krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”

Gilbert Ryle

(1900-1976)



krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
 - ▶ *Czym jest umysł?*, 1970 (*The Concept of Mind*, 1949)
 - ▶ dualizm jako błąd kategorialny

krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową

krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową
 - ▶ przyczynowość niefizyczalna – zasada kauzalnego domknięcia dziedziny fizycznej
 - ▶ zasada zachowania energii
 - ▶ brak anomalii w pracy mózgu

krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową
- ▶ zarzuty metanaukowe
 - ▶ jałowość eksplanacyjna dualizmu
 - ▶ brzytwa Ockhama

William Ockham

(1287-1347)



krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową
- ▶ zarzuty metanaukowe
 - ▶ jałowość eksplanacyjna dualizmu
 - ▶ brzytwa Ockhama
 - ▶ skuteczność badań w paradygmacie fizykalistycznym (=monistycznym)

krytyka dualizmu substancji

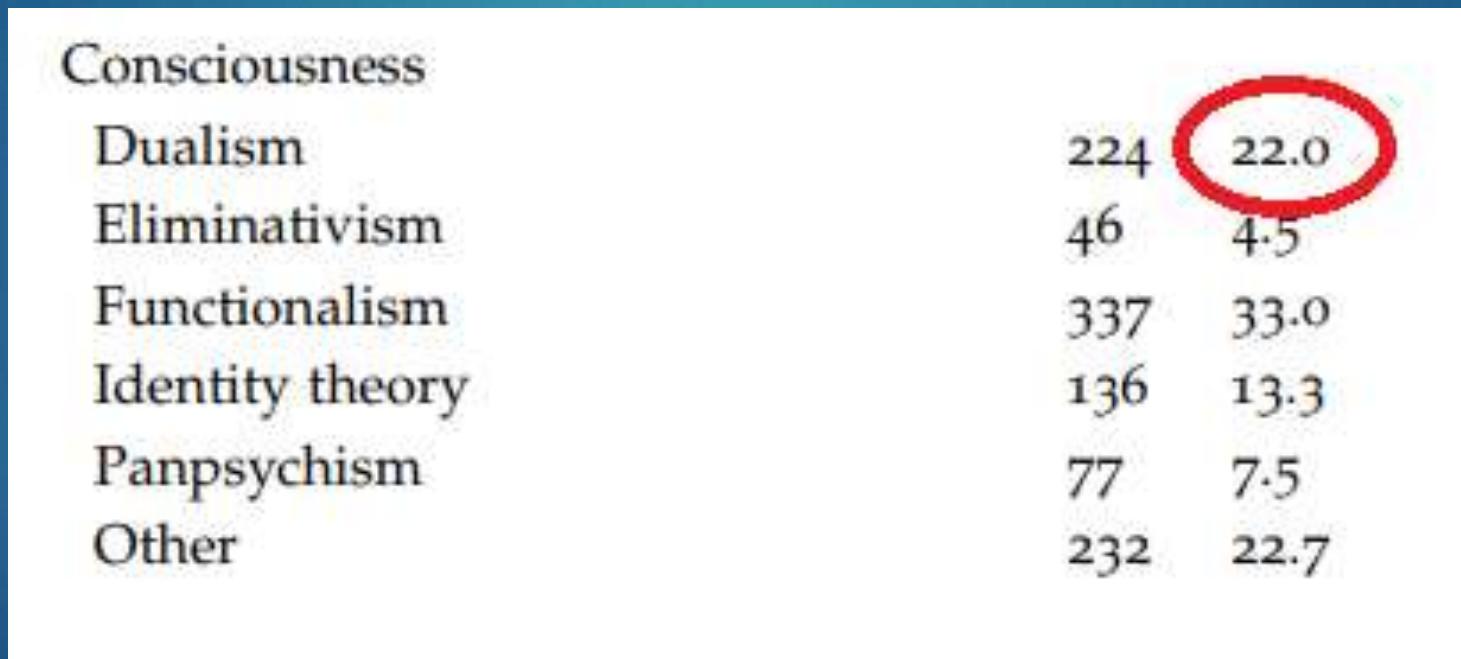
- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową
- ▶ zarzuty metanaukowe
- ▶ myślenie jako istota duszy
 - ▶ co się dzieje z duszą, gdy nie myślimy?
 - ▶ co z procesami nieświadomymi?

krytyka dualizmu substancji

- ▶ „mit ducha w maszynie”
- ▶ niespójność z wiedzą naukową
- ▶ zarzuty metanaukowe
- ▶ myślenie jako istota duszy
- ▶ Daniel Dennett: dualizm jako kapitulacja

czy dualizm substancji jest do obrony?

David Bourget, David Chalmers, *Philosophers on Philosophy: The 2020 Philpapers Survey*, „Philosophers' Imprint” 23 (1), 2023



czy dualizm substancji jest do obrony?

► współczesny argument za dualizmem – Alvin Plantinga, Richard Swinburne

P₁ Mogę istnieć bez mojego ciała

P₂ Moje ciało nie może istnieć bez mojego ciała

P₃ Nie jestem identyczny z moim ciałem

P₄ Jestem z istoty albo ciałem, albo substancją niefizyczną

W Jestem z istoty substancją niefizyczną



wstęp do filozofii umysłu i sztucznej inteligencji

19.10.2023

WYKŁAD 3

czy dualizm substancji jest do obrony?

► współczesny argument za dualizmem – Alvin Plantinga, Richard Swinburne

P₁ Mogę istnieć bez mojego ciała

P₂ Moje ciało nie może istnieć bez mojego ciała

P₃ Nie jestem identyczny z moim ciałem

P₄ Jestem z istoty albo ciałem, albo substancją niefizyczną

W Jestem z istoty substancją niefizyczną

- ▶ jak badać umysł?
 - ▶ uprzywilejowana perspektywa pierwszoosobowa
 - ▶ introspekcja

Wilhelm Wundt
(1832-1920)



- ▶ jak badać umysł?
 - ▶ uprzywilejowana perspektywa pierwszoosobowa
 - ▶ introspekcja
 - ▶ badanie naukowe powinno mieć charakter intersubiektywny

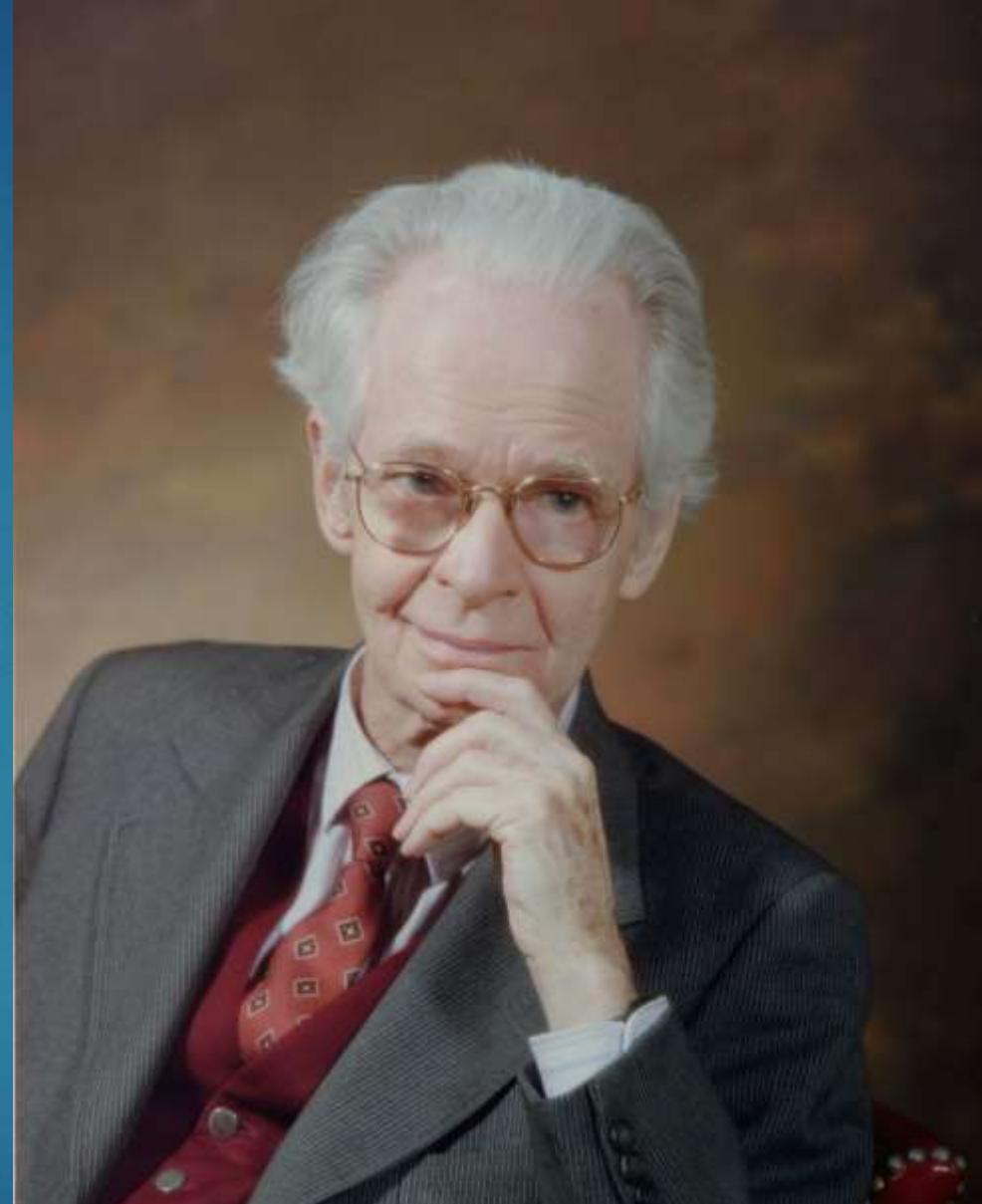
behawioryzm



John Broadus Watson
(1878-1958)



Burrhus Frederick Skinner
(1904-1990)



behawioryzm

- ▶ teoria psychologiczna
 - ▶ mechanizm bodziec-reakcja
 - ▶ badanie zachowania
 - ▶ behawioryzm metodologiczny
 - ▶ odrzucenie języka stanów mentalnych

Gilbert Ryle

(1900-1976)



behawioryzm

- ▶ teoria filozoficzna
 - ▶ behawioryzm logiczny/analytyczny: stany umysłu są tożsame z zachowaniem bądź dyspozycją do zachowania
 - ▶ analizy dyspozycyjne
 - ▶ argument za behawioryzmem:

P₁. Zjawiska umysłowe nie są dostępne intersubiektywnie

P₂. Odnoszące się do nich wyrażenia językowe wtedy należą do intersubiektywnego języka, gdy istnieje kryterium ich użycia

P₃. Nie ma innego kryterium niż odnoszące się do zachowania (dyspozycji)

W. Zjawiska umysłowe tożsame są z zachowaniem/dyspozycją

behawioryzm

- ▶ teoria filozoficzna
- ▶ krytyka behawioryzmu
 - 1. Te same bodźce nie skutkują tym samym zachowaniem
 - 2. Dyspozycje do zachowań nie są (logicznie) tożsame ze stanami mentalnymi
 - 3. Kolistość wyjaśnień behawiorystycznych („Jan jest głodny”...)
 - 4. Niewytłumaczalność innych (obliczenia w głowie)
 - 5. Kwestia snu

teoria identyczności



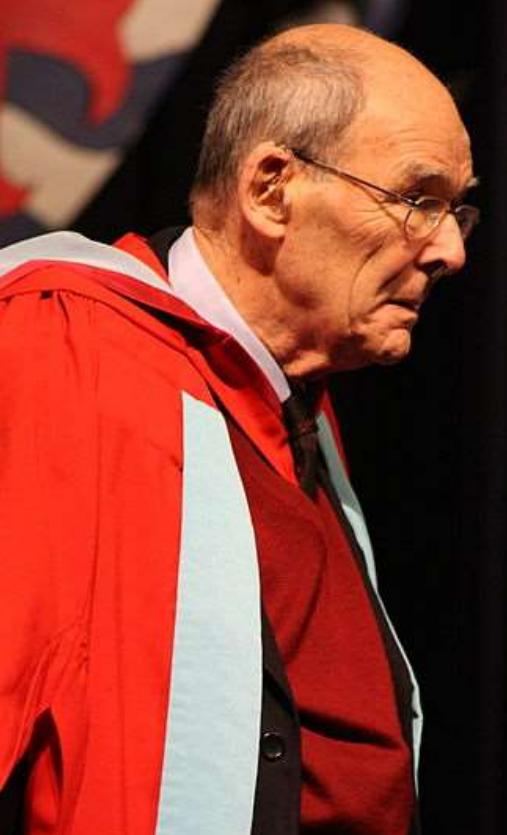


John Jamieson Carswell Smart
(1920-2012)

Ullin Thomas Place
(1924-2000)



David Malet Armstrong
(1926-2014)



Did this brain contain the consciousness of U.T. Place?

teoria identyczności

Nasze rozważania nad behawioryzmem nauczyły nas dwu rzeczy: 1. nie można zaprzeczyć istnienia wewnętrznych stanów umysłowych, 2. mimo to zachowanie zewnętrzne jest w jakiś sposób włączone w pojęcie umysłu.

David M. Armstrong, *Materialistyczna teoria umysłu*, PWN 1982

teoria identyczności

- ▶ umysł = mózg
- ▶ umysł jest identyczny z mózgiem
- ▶ identyczność jako rodzaj redukcji
 - ▶ sprowadzenie jednych właściwości do drugich



teoria identyczności

- ▶ umysł = mózg
- ▶ umysł jest identyczny z mózgiem
- ▶ identyczność jako rodzaj redukcji
 - ▶ sprowadzenie jednych własności do drugich
 - ▶ błyskawica = wyładowanie elektryczne (nie np. gniew boga, żywioł...)

teoria identyczności

- ▶ umysł = mózg
- ▶ umysł jest identyczny z mózgiem
- ▶ identyczność jako rodzaj redukcji
 - ▶ sprowadzenie jednych własności do drugich
 - ▶ błyskawica = wyładowanie elektryczne
 - ▶ redukcja ontologiczna – między elementami świata (przedmiotami, własnościami, zdarzeniami, procesami...)
 - ▶ redukcja reprezentacyjna – między elementami reprezentującymi świat (pojęciami, teoriami, modelami, schematami...)

teoria identyczności

- ▶ umysł = mózg
- ▶ umysł jest identyczny z mózgiem
- ▶ identyczność jako rodzaj redukcji
- ▶ identyczność to coś więcej niż twierdzenie, że stany umysłowe korelują ze stanami fizycznymi mózgu
 - ▶ spójka „jest” wyraża ścisłą identyczność
 - ▶ „stan umysłu” i „stan mózgu” mają różne znaczenie, ale to samo odniesienie
 - ▶ logiki języków mentalnego i materialnego są różne

teoria identyczności

- ▶ umysł = mózg
- ▶ umysł jest identyczny z mózgiem
- ▶ identyczność jako rodzaj redukcji
- ▶ identyczność to coś więcej niż twierdzenie, że stany umysłowe korelują ze stanami fizycznymi mózgu
- ▶ Place: teoria identyczności może zostać potwierdzona przez naukę
- ▶ Smart: to teoria metafizyczna, nierożstrzygalna naukowo
 - ▶ wszystkie fakty empiryczne zgodne z dualizmem i teorią identyczności
 - ▶ ale dualizm postuluje istnienie dziwacznych praw psychofizycznych
 - ▶ zatem zgodnie z brzytwą Ockhama powinniśmy wybrać teorię identyczności

identyczność typów i egzemplarzy

- ▶ typy i egzemplarze



pięć egzemplarzy jednego typu (kot)



KOT

KOT

KOT

KOT

KOT

pięć egzemplarzy jednego typu (słowa KOT)

identyczność typów i egzemplarzy

- ▶ typy i egzemplarze
- ▶ teoria identyczności typów (*type-type identity*)
 - ▶ u każdego człowieka aktywność włókien nerwowych C jest identyczna z bólem typu K
 - ▶ wyobrażenie Mostu Grunwaldzkiego jest identyczne z aktywnością włókien nerwowych XYZ
 - ▶ teza o wielorakiej realizowalności
 - ▶ Armstrong: „materializm praw dodatkowych” → dowolny system o wystarczającej złożoności

identyczność typów i egzemplarzy

- ▶ typy i egzemplarze
- ▶ teoria identyczności typów (*type-type identity*)
- ▶ teoria identyczności egzemplarzy (*token-token identity*)
 - ▶ u człowieka X aktywność włókien nerwowych C jest identyczna z bólem typu K
 - ▶ u człowieka X wyobrażenie Mostu Grunwaldzkiego jest identyczne z aktywnością włókien nerwowych XYZ

krytyka teorii identyczności

- ▶ jakości procesów umysłowych i procesów mózgowych są różne
 - ▶ ból może być intensywny, punktowy, rwący, tępy..., ale aktywność włókien C taka nie jest
 - ▶ aktywność włókien C może mieć określone (w miliwoltach) liczbowo napięcie, ale ból tej własności nie ma

krytyka teorii identyczności

- ▶ jakości procesów umysłowych i procesów mózgowych są różne
- ▶ *qualia*

wstęp do filozofii umysłu i sztucznej inteligencji

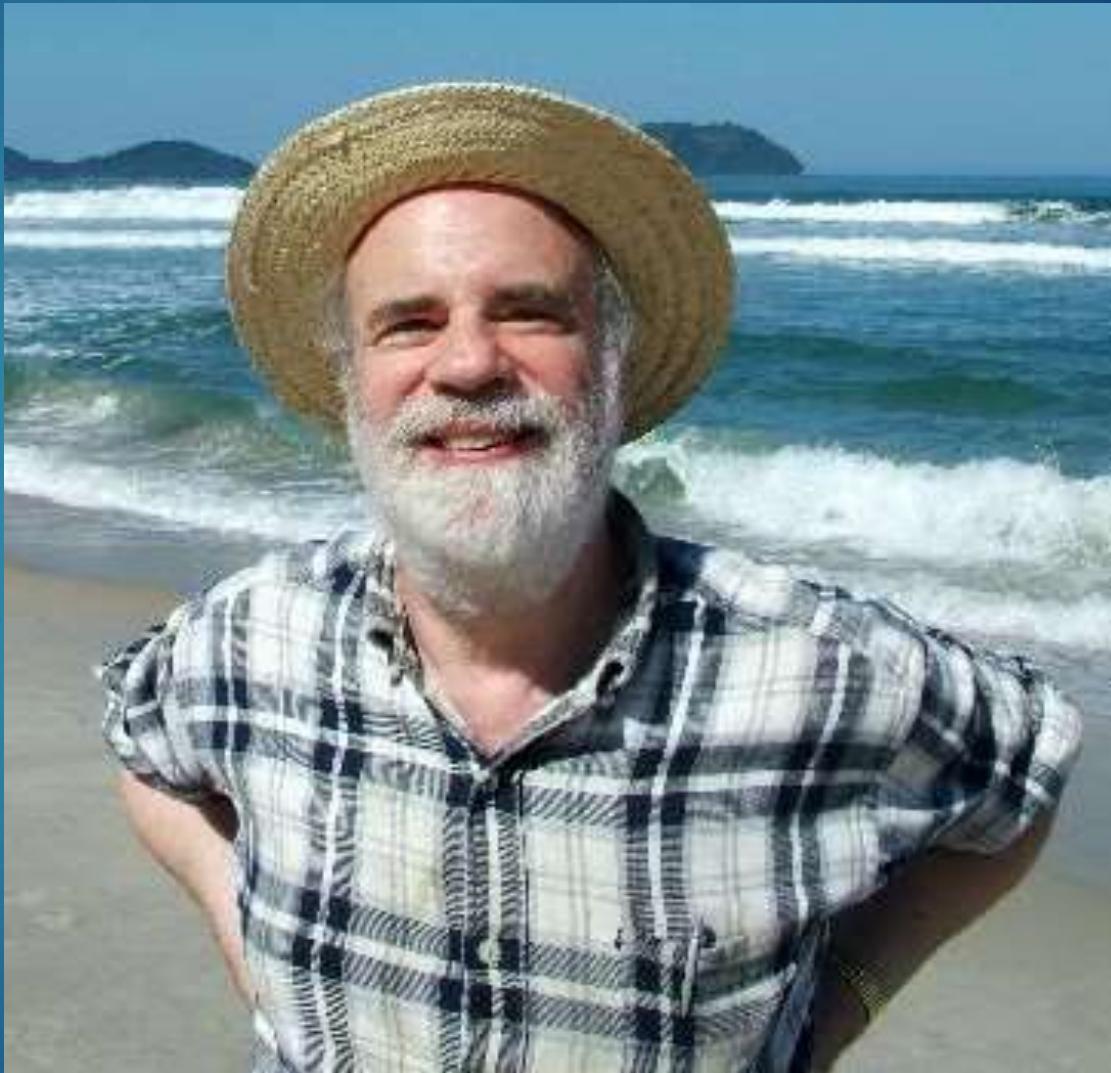
26.10.2023

WYKŁAD 4

krytyka teorii identyczności

- ▶ jakości procesów umysłowych i procesów mózgowych są różne
- ▶ *qualia*
- ▶ Saul Kripke – sztywny desygnator

Saul Kripke
(1940-2022)



eliminatywizm

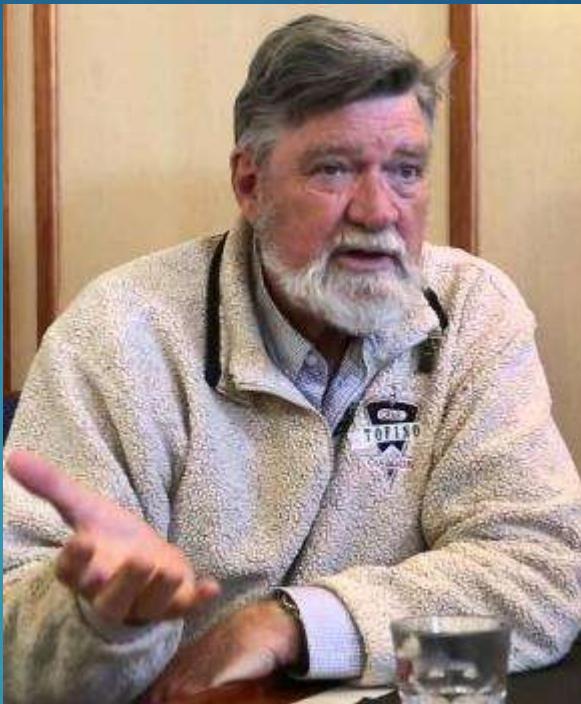




Paul Churchland
(ur. 1942)



Patricia Churchland
(ur. 1943)



Stephen Stich
(ur. 1943)

eliminatywizm

- ▶ pojęcie sądu
- ▶ sąd p
 - ▶ *proposition*

eliminatywizm

- ▶ pojęcie sądu
- ▶ sąd p
- ▶ p – treść sądu
- ▶ nastawienie sądzeniowe (*propositional attitude*)
 - ▶ sądzi, że p
 - ▶ myśli, że p
 - ▶ uważa, że p
 - ▶ jest przekonany, że p
 - ▶ wątpi, że p

eliminatywizm

- ▶ pojęcia mentalne nie z pierwszoosobowego dostępu, ale z języka potocznego
- ▶ psychologia potoczna (*folk psychology*) jako teoria
 - ▶ przewidywania
 - ▶ wyjaśnienia
 - ▶ ścisłe prawa

eliminatywizm

- ▶ pojęcia mentalne nie z pierwszoosobowego dostępu, ale z języka potocznego
- ▶ psychologia potoczna (*folk psychology*) jako teoria
- ▶ redukcja eliminacyjna

eliminatywizm

- P₁ Psychologia potoczna jest teorią empiryczną (ze wszystkimi tego konsekwencjami)
- P₂ Teorie empiryczne należy oceniać pod względem właściwości empirycznych i pragmatycznych
- P₃ Ocena psychologii potoczej nie wypada korzystnie
- P₄ Istnieje lepsza pod tym względem teoria konkurencyjna
- W Należy wyeliminować psychologię potoczną i przyjąć lepszą teorię

eliminatywizm

- P₁ Psychologia potoczna jest teorią empiryczną (ze wszystkimi tego konsekwencjami)
- P₂ Teorie empiryczne należy oceniać pod względem właściwości empirycznych i pragmatycznych
- P₃ Ocena psychologii potoczej nie wypada korzystnie
 1. Rażąca nieskuteczność eksplanacyjna
 2. Zero postępu przez całe wieki
 3. Mało prawdopodobne, że zostanie częścią obrazu naukowego
- P₄ Istnieje lepsza pod tym względem teoria konkurencyjna
- W Należy wyeliminować psychologię potoczną i przyjąć lepszą teorię

eliminatywizm

- ▶ pojęcia mentalne nie z pierwszoosobowego dostępu, ale z języka potocznego
- ▶ psychologia potoczna (*folk psychology*) jako teoria
- ▶ redukcja eliminacyjna
 - ▶ usunięcie określonych obiektów z uniwersum dyskursu
 - ▶ tak jak usunęliśmy z dyskursu naukowego: flogiston (cieplik), eter, pneumę, czarownice etc.
 - ▶ z dyskursu o umyśle usuniemy w końcu: qualia, przekonania, nastawienia sądzeniowe, świadomość etc.
 - ▶ psychologia potoczna zostanie zastąpiona lepszą, fizykalistyczną teorią neurofizjologiczną

eliminatywizm

- ▶ pojęcia mentalne nie z pierwszoosobowego dostępu, ale z języka potocznego
- ▶ psychologia potoczna (*folk psychology*) jako teoria
- ▶ redukcja eliminacyjna
- ▶ eliminatywizm analityczny (Stich)
 - ▶ psychologia potoczna nie jest słabszą teorią, ale niemożliwą do zrealizowania
 - ▶ kłopot z pojęciem *treści* sądu oraz jego atrybucji
 - ▶ psychologia potoczna musi zostać zastąpiona przez dojrzałą kognitywistykę

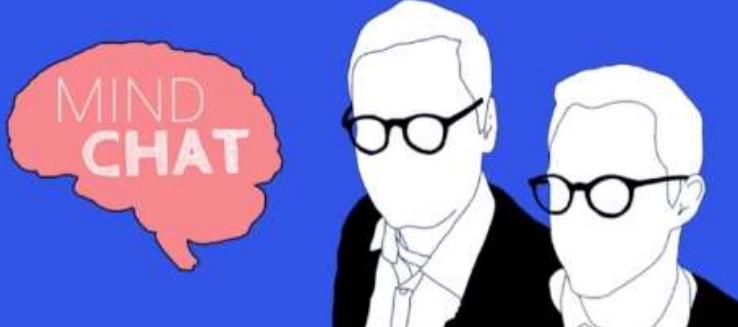
iluzjonizm



The picture can't be displayed.

Keith Frankish
(ur. 1962)

MIND CHAT GOFF & FRANKISH



Mind Chat

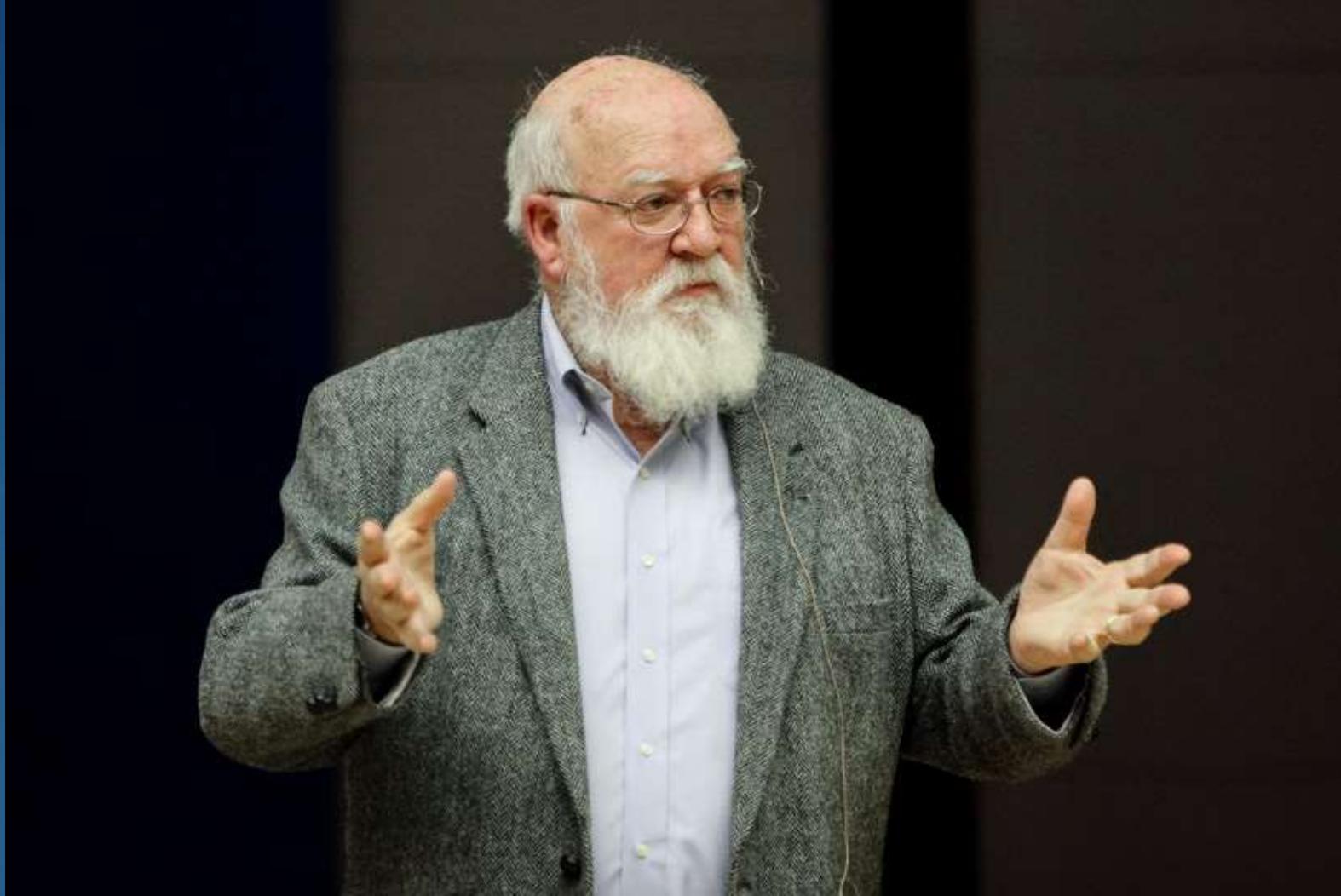
@MindChat · 6,28 tys. subskrybentów · 31 filmów

Philip Goff is a philosopher who thinks consciousness pervades the universe. Keith Frankis... >

philipgoffphilosophy.com i jeszcze 2 linki



Subskrybujesz



Daniel Dennett
(ur. 1942)

iluzjonizm

- ▶ forma eliminatywizmu, ale głównie w odniesieniu do świadomości
- ▶ świadomość jako iluzja
 - ▶ złudzenie, sztuczka naszego mózgu
 - ▶ najlepsza strategia eksplanacyjna
 - ▶ mechanizm adaptacyjny
 - ▶ podobnie jak złudzenie jaźni → Dennett: środek narracyjnej ciężkości

iluzjonizm

- ▶ forma eliminatywizmu, ale głównie w odniesieniu do świadomości
- ▶ świadomość jako iluzja
 - ▶ złudzenie, sztuczka naszego mózgu
 - ▶ najlepsza strategia eksplanacyjna
 - ▶ mechanizm adaptacyjny
 - ▶ podobnie jak złudzenie jaźni
 - ▶ Dennett: filozof jak iluzjonista – na scenie dzieje się *magia*, filozof wyjaśnia, co tak naprawdę się za tym kryje

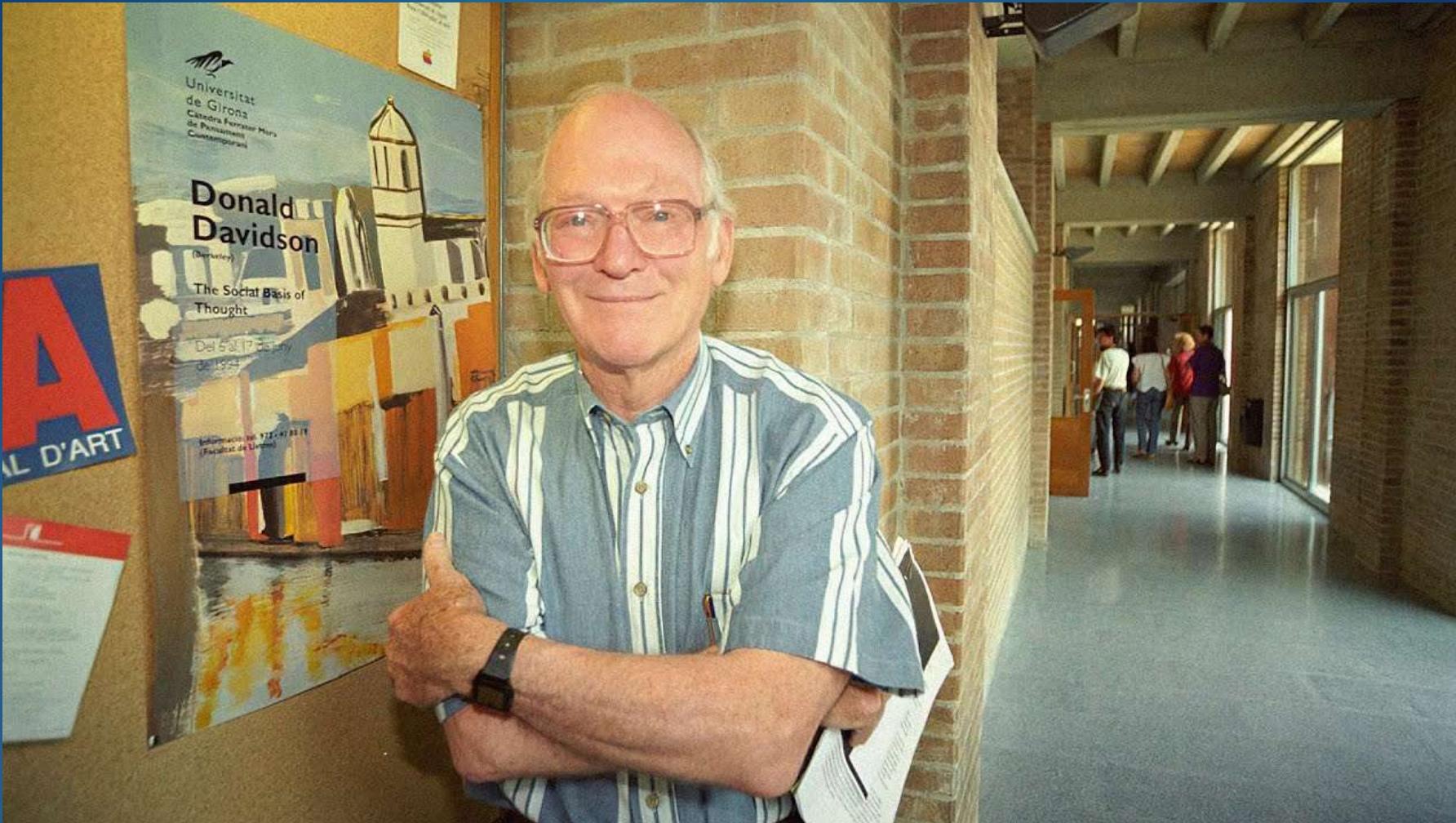
wstęp do filozofii umysłu i sztucznej inteligencji

9.11.2023

WYKŁAD 5

monizm anomalny





Donald Davidson
(1917-2003)

monizm anomalny

- ▶ materializm niereductyjny

monizm anomalny

- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
 - ▶ monizm nomologiczny – np. teoria identyczności
 - ▶ dualizm nomologiczny – np. interakcjonizm, paralelizm, epifenomenalizm
 - ▶ dualizm anomalny – np. kartezjanizm
 - ▶ monizm anomalny

monizm anomalny

- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
- ▶ zdarzenia mentalne wyróżnia nie prywatność, subiektywność, niematerialność – ale intencjonalność

monizm anomalny

- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
- ▶ zdarzenia mentalne wyróżnia nie prywatność, subiektywność, niematerialność – ale intencjonalność
- ▶ wszystkie zdarzenia są fizyczne, ale nie ma praw łączących umysł z mózgiem
 - ▶ pojęcia mentalne nie redukują się do fizycznych
 - ▶ odróżnienie sfery ontologicznej (redukacja) od językowej (brak redukcji)
 - ▶ stany mentalne od fizycznych różni sposób opisu – język stanów mentalnych jest autonomiczny
 - ▶ pojęcia mentalne różnych ludzi nie są równoważne

Most
Grunwaldzki
jest piękny

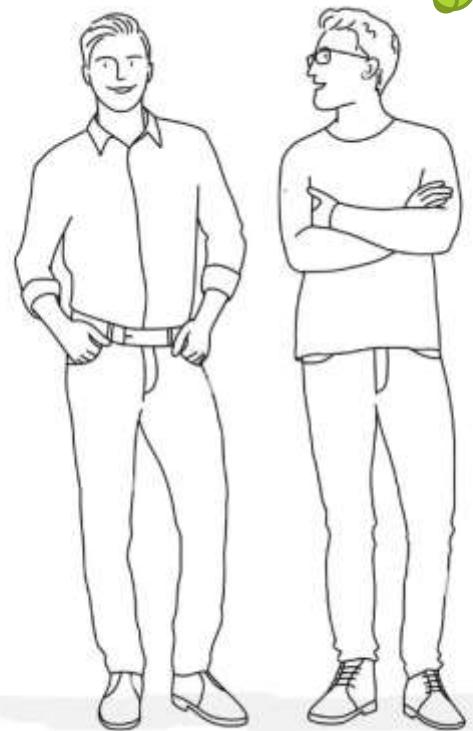
Most
Grunwaldzki
jest piękny





Grunwald
Bridge is
beautiful

Grunwald
Bridge are
beautiful



monizm anomalny

- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
- ▶ zdarzenia mentalne wyróżnia nie prywatność, subiektywność, niematerialność – ale intencjonalność
- ▶ wszystkie zdarzenia są fizyczne, ale nie ma praw łączących umysł z mózgiem
 - ▶ pojęcia mentalne nie redukują się do fizycznych
 - ▶ odróżnienie sfery ontologicznej (redukacja) od językowej (brak redukcji)
 - ▶ stany mentalne od fizycznych różni sposób opisu – język stanów mentalnych jest autonomiczny
 - ▶ pojęcia mentalne różnych ludzi nie są równoważne
 - ▶ dualizm na poziomie języka

monizm anomalny

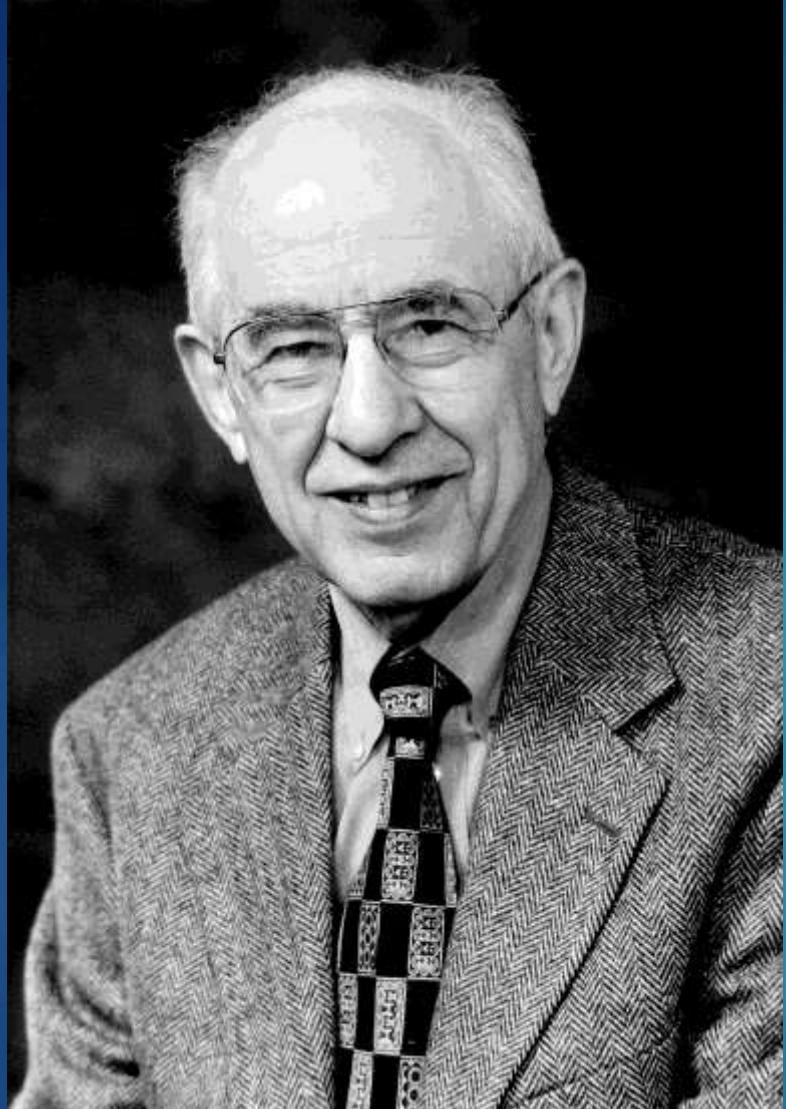
- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
- ▶ zdarzenia mentalne wyróżnia nie prywatność, subiektywność, niematerialność – ale intencjonalność
- ▶ wszystkie zdarzenia są fizyczne, ale nie ma praw łączących umysł z mózgiem
- ▶ superweniencja tego, co mentalne, na tym, co fizyczne

monizm anomalny

- ▶ materializm niereductyjny
- ▶ krzyżujący się podział stanowisk w kwestii natury i przyczynowości umysłu
- ▶ zdarzenia mentalne wyróżnia nie prywatność, subiektywność, niematerialność – ale intencjonalność
- ▶ wszystkie zdarzenia są fizyczne, ale nie ma praw łączących umysł z mózgiem
- ▶ **superweniencja** tego, co mentalne, na tym, co fizyczne
 - ▶ A superweniuje na B, wówczas konieczne jest zajście jakieś zmiany w B, by możliwe było zajście zmiany w A
 - ▶ własności bazowe (podstawowe) i superwenientne (nadbudowane)
 - ▶ trzy postulaty: kowariancja, zależność, niereductowalność

funkcjonalizm





Hilary Putnam
(1926-2016)



Jerry Fodor
(1935-2017)

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
 - ▶ teoria identyczności: krzesło jest *identyczne* z pewnym określonym drewnianym przedmiotem



funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
 - ▶ teoria identyczności: krzesło jest *identyczne* z pewnym określonym drewnianym przedmiotem
 - ▶ funkcjonalizm: krzesło to przedmiot służący do siedzenia, bez względu na materiał, z jakiego jest wykonane



funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
 - ▶ teoria identyczności: krzesło jest identyczne z pewnym określonym drewnianym przedmiotem
 - ▶ funkcjonalizm: krzesło to przedmiot *służący do siedzenia, bez względu na materiał, z jakiego jest wykonane*

funkcjonalizm

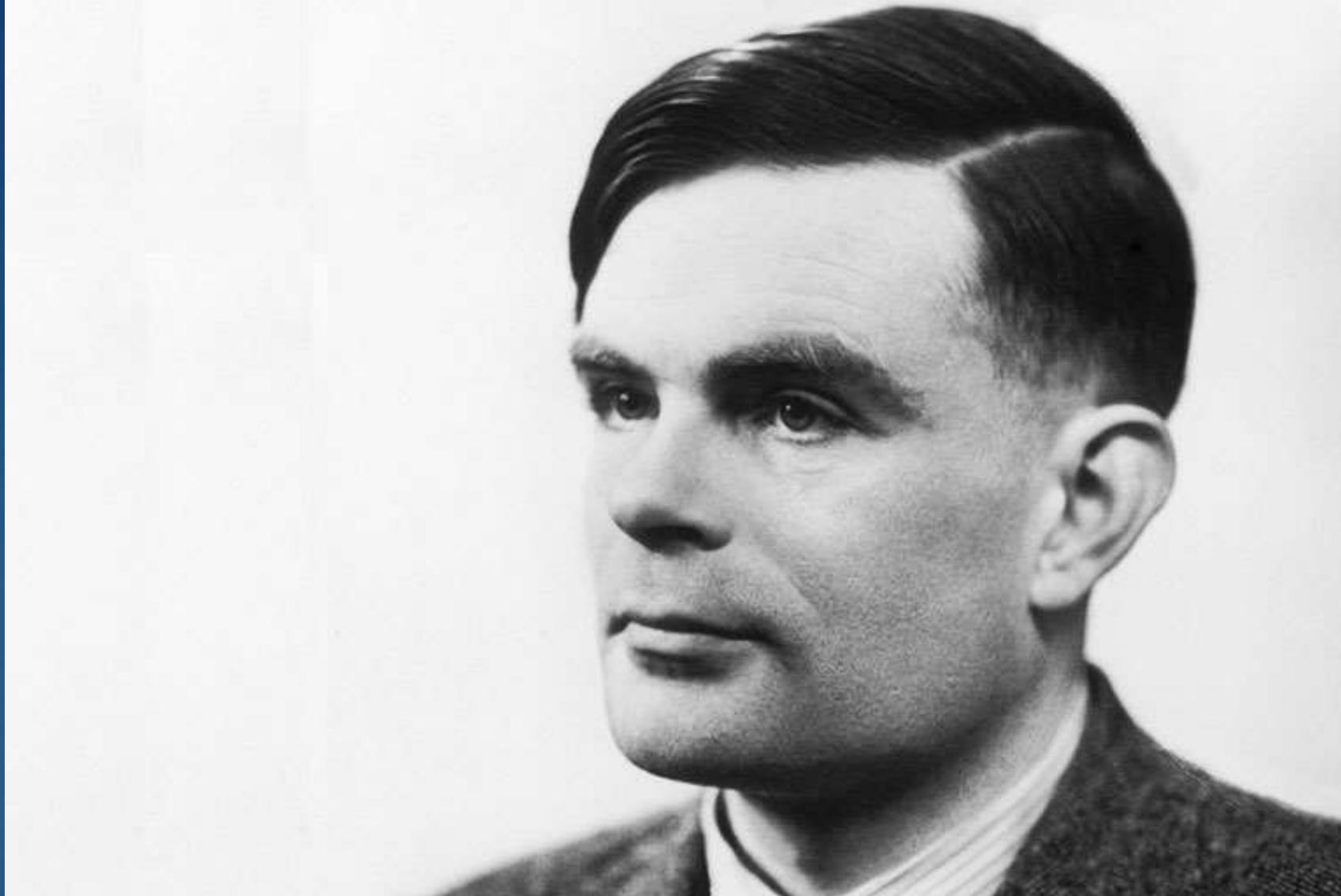
- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
 - ▶ stan funkcjonalny całego organizmu
 - ▶ np. stan bólu – funkcjonalny stan organizmu rozumianego jako odpowiedni system materialny

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
- ▶ schemat bodziec–reakcja
 - ▶ *input–output*
 - ▶ behawioryzm+
 - ▶ stany mentalne to funkcja między wejściem a wyjściem
 - ▶ zamiast „dyspozycji” – „role przyczynowe/funkcjonalne”

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
- ▶ schemat bodziec–reakcja
- ▶ komputacionizm



Alan Turing
(1912-1954)

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
- ▶ schemat bodziec–reakcja
- ▶ komputacionizm
 - ▶ maszyna Turinga
 - ▶ stany umysłu to stany obliczeniowe
 - ▶ człowiek jako pewna forma komputera
 - ▶ mózg i umysł mają się do siebie jak *hardware* i *software*
 - ▶ umysł jako algorytm, program realizowany na mózgu
 - ▶ kluczowe założenie w programie badawczym sztucznej inteligencji

funkcjonalizm

„hipoteza funkcjonalna” Putnama (przykład bólu) → hipoteza empiryczna

1. Wszystkie organizmy zdolne do odczuwania bólu są automatami probabilistycznymi
2. Bycie zdolnym do odczuwania bólu to posiadanie odpowiedniej organizacji funkcjonalnej
3. Żaden organizm zdolny do odczuwania bólu nie dzieli się na części posiadające część organizacji funkcjonalnej
4. W przypadku każdego takiego automatu probabilistycznego istnieje podzbiór danych wejściowych taki, że organizm odczuwa ból wówczas gdy niektóre z jego zmysłowych danych wejściowych należą do tego zbioru

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
- ▶ schemat bodziec–reakcja
- ▶ komputacionizm
- ▶ neutralność wobec rozstrzygnięć metafizycznych (dualizm, monizm)
- ▶ teza o wielorakiej realizowalności

wstęp do filozofii umysłu i sztucznej inteligencji

16.11.2023

WYKŁAD 6

funkcjonalizm

- ▶ znaczenie ma funkcja, nie materia
- ▶ stan mentalny = stan funkcjonalny
- ▶ schemat bodziec–reakcja
- ▶ komputacionizm
- ▶ neutralność wobec rozstrzygnięć metafizycznych (dualizm, monizm)
- ▶ teza o wielorakiej realizowalności

krytyka funkcjonalizmu



Ned Block
(ur. 1942)



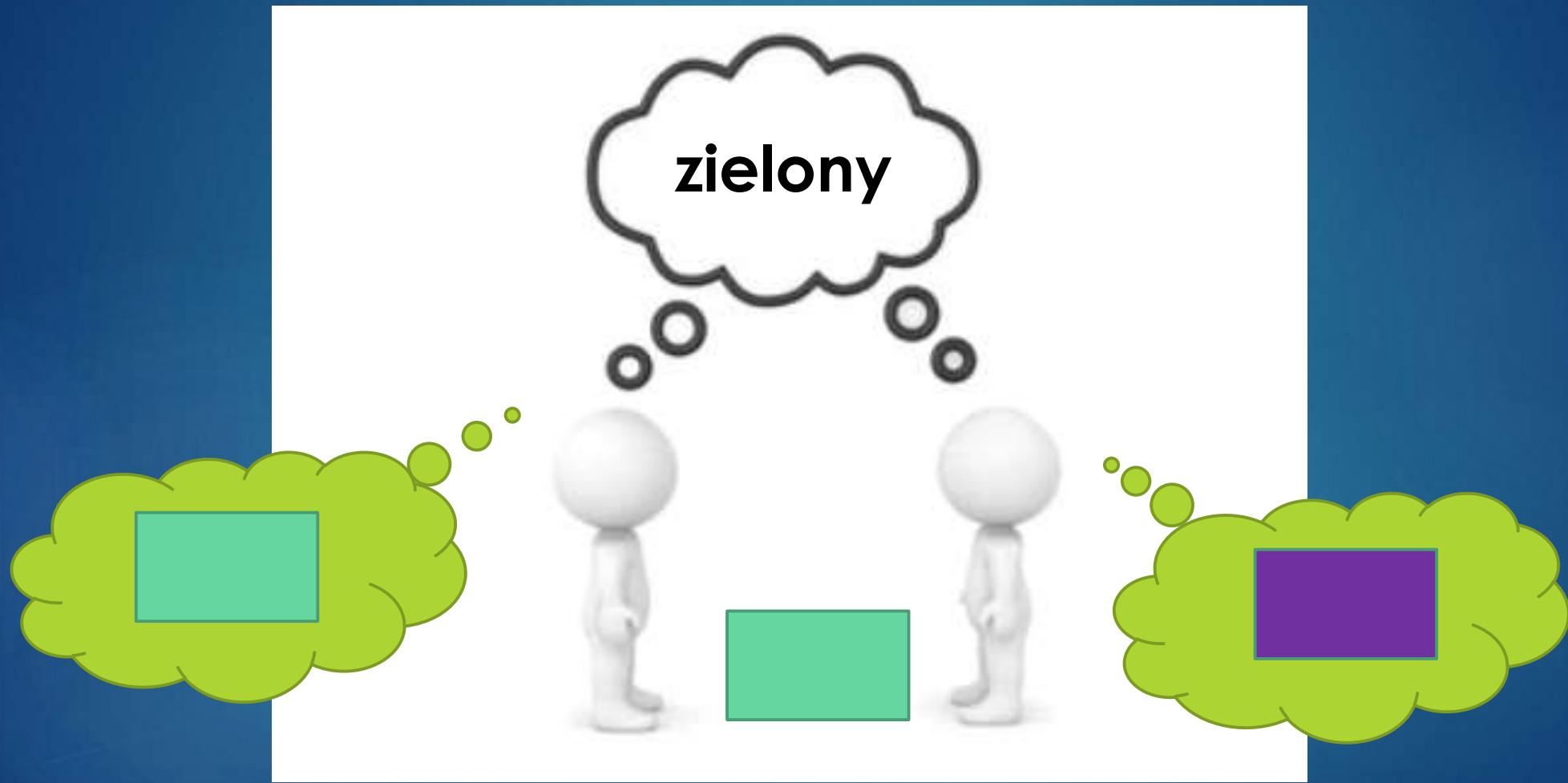
krytyka funkcjonalizmu

- ▶ argument „chińskiego narodu” (eksperyment myślowy)



krytyka funkcjonalizmu

- ▶ argument „chińskiego narodu” (eksperyment myślowy)
- ▶ odwrócone *qualia*



koneksjonizm



koneksjonizm

- ▶ założenia jak w komputacjonizmie (?)

koneksjonizm

- ▶ założenia jak w komputacjonizmie
 - ▶ mózg jak komputer
 - ▶ stany mentalne to stany obliczeniowe

Donald Hebb
(1904-1985)

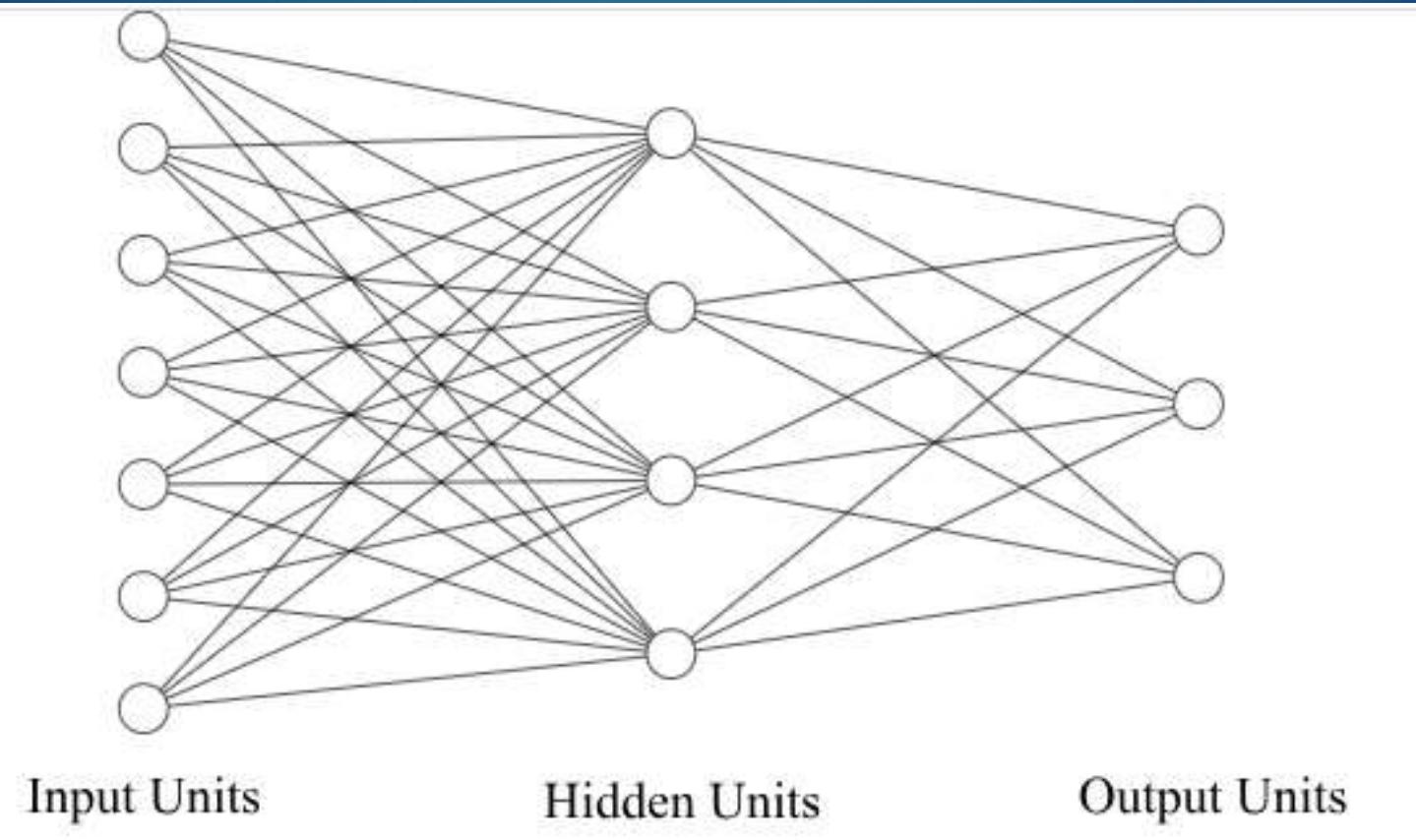


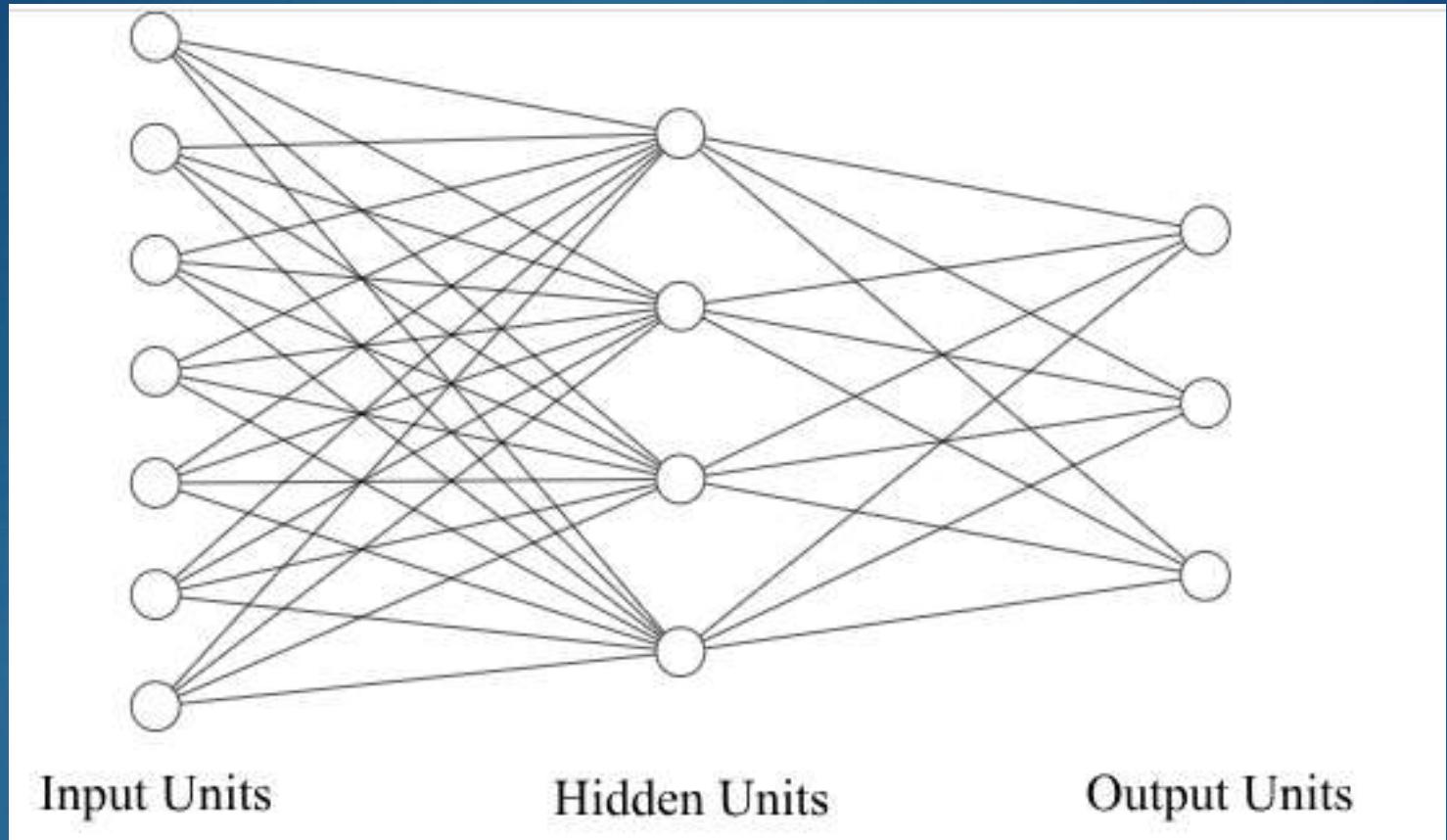
koneksjonizm

- ▶ założenia jak w komputacjonizmie
- ▶ paradygmat popularny w kognitywistyce

koneksjonizm

- ▶ założenia jak w komputacjonalizmie
- ▶ paradymat popularny w kognitywistyce
- ▶ modelowanie umysłu
 - ▶ budowa i funkcjonowanie sieci neuronowych
 - ▶ sztuczne sieci neuronowe
 - ▶ przetwarzanie informacji w sposób równoległy, a nie sekwencyjny





Input Units

Hidden Units

Output Units

neurony czuciowe
(zmysłowe)

pozostałe neurony

neurony ruchowe

koneksjonizm

- ▶ założenia jak w komputacjonizmie
- ▶ paradygmat popularny w kognitywistyce
- ▶ modelowanie umysłu
 - ▶ budowa i funkcjonowanie sieci neuronowych
 - ▶ sztuczne sieci neuronowe
 - ▶ *brain style modelling*

koneksjonizm

- ▶ założenia jak w komputacjonalizmie
- ▶ paradygmat popularny w kognitywistyce
- ▶ modelowanie umysłu
- ▶ sieciowe uczenie się
 - ▶ zamiast programowania algorytmicznego
 - ▶ trening, wzmacnianie

koneksjonizm

- ▶ założenia jak w komputacjonalizmie
- ▶ paradymat popularny w kognitywistyce
- ▶ modelowanie umysłu
- ▶ sieciowane uczenie się
- ▶ pamięć asocjacyjna (skojarzeniowa)
 - ▶ z szumów na wejściu → „oczyszczone” schematy pojęciowe na wyjściu
 - ▶ mniejszy wpływ uszkodzeń

modularyzm





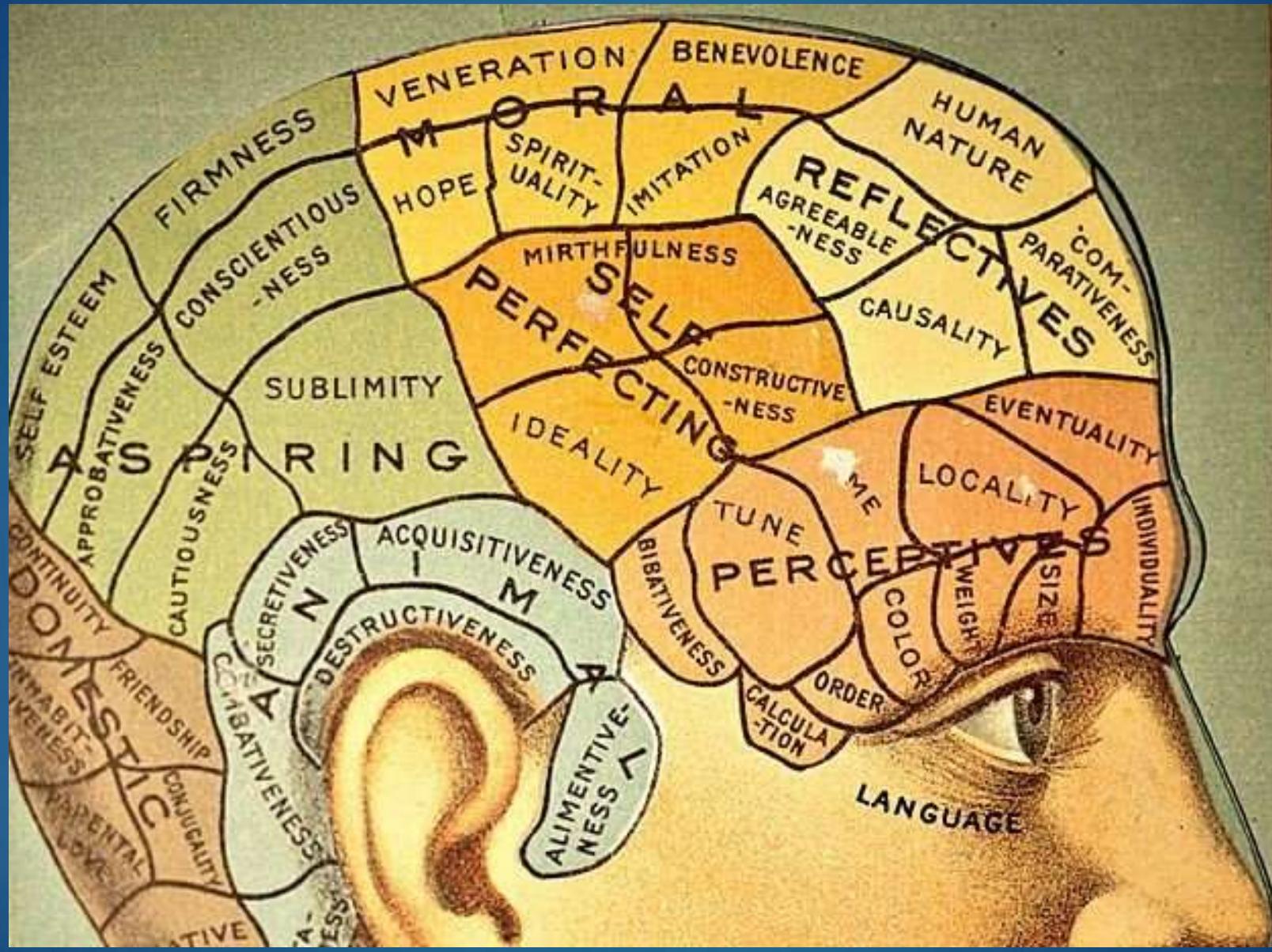
Jerry Fodor
(1935-2017)



modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta





Franz Joseph Gall
(1758-1828)



DOZORCA

do Doktora

Wać pan przyszedłeś zwiedzić szpital waryjatów?

(...)

Cały szpital rozłożę jak zegar po sztuce...

Pan zapewne w medycznej ćwiczysz się nauce?

DOKTOR

Tak.

DOZORCA

Jakież systemata pan doktor zachwala?

DOKTOR

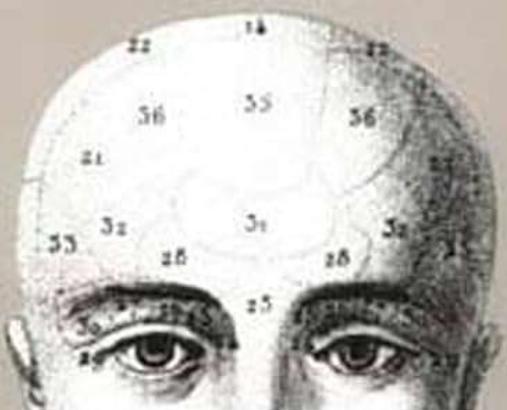
Macanie głów...

DOZORCA

Rozumiem, to systemat Galla.

THE MODULARITY OF MIND

Jerry A. Fodor



modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta
- ▶ moduły mają charakter autonomiczny
 - ▶ hermetyczność informacyjna (brak przepływu między modułami)
 - ▶ niedostępność dla centralnych przekonań (nie zmieniają przepływu informacji w modułach)
 - ▶ specjalizacja (każdy moduł przeznaczony do rozwiązywania skomplikowanego problemu poznawczego)
 - ▶ wrodzoność (przypisane modułom informacje są zaprogramowane genetycznie)
 - ▶ automatyzm (funkcjonowanie modułów nie podlega kontroli)

modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta
- ▶ moduły mają charakter autonomiczny
- ▶ dwa rodzaje podsystemów
 - ▶ moduły i przekaźniki
 - ▶ procesory centralne – często mają charakter nieświadomy, testują hipotezy o świecie

modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta
- ▶ moduły mają charakter autonomiczny
- ▶ dwa rodzaje podsystemów
- ▶ założenia modularyzmu leżą u podstaw programu neuropsychologii poznawczej

Michael Gazzaniga
(ur. 1939)



modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta
- ▶ moduły mają charakter autonomiczny
- ▶ dwa rodzaje podsystemów
- ▶ założenia modularitymu leżą u podstaw programu neuropsychologii poznawczej
 - ▶ aktywność umysłu jako rezultat skoordynowanego działania modułów
 - ▶ zakłócenia poznawcze związane z uszkodzeniem mózgu dają podstawę do identyfikacji wadliwego modułu

modularyzm

- ▶ właściwa architektura umysłu ma charakter modułowy
- ▶ frenologia jako protoplasta
- ▶ moduły mają charakter autonomiczny
- ▶ dwa rodzaje podsystemów
- ▶ założenia modularyzmu leżą u podstaw programu neuropsychologii poznawczej
- ▶ filozoficzna konkluzja modularyzmu: nie ma czegoś takiego jak umysł
 - ▶ ...o ile rozumiemy przez to jedną określoną rzecz czy substancję
 - ▶ umysł to nazwa zbiorcza na funkcje, stany czy własności realizowane przez poszczególne wyspecjalizowane moduły
 - ▶ zdaniem Fodora nie ma jedności świadomości

Michael S.
Gazzaniga

Instynkt świadadomości

Jak z mózgu
wyłania się umysł?

Ożywca i poszerzająca horyzonty opowieść
o naukowych i filozoficznych aspektach świadomości.
Michael S. Gazzaniga objaśnia niektóre z najtrudniejszych pytań,
jakie potrafi sformułować ludzki umysł.

– Steven Pinker



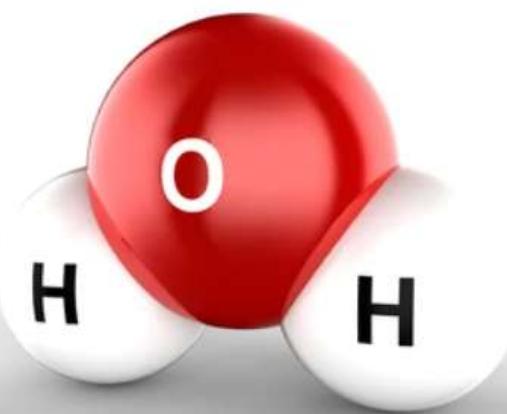
wstęp do filozofii umysłu i sztucznej inteligencji

23.11.2023

WYKŁAD 7

emergentyzm

- ▶ **emergo (emergere)** – wyłaniam się, powstaję
- ▶ wyłonienie się jakościowo nowych własności wyższego rzędu z oddziaływań na poziomie niższym
- ▶ „całość jest czymś więcej niż sumą części”
 - ▶ nowe własności w agregatach, nieobecne w częściach składowych



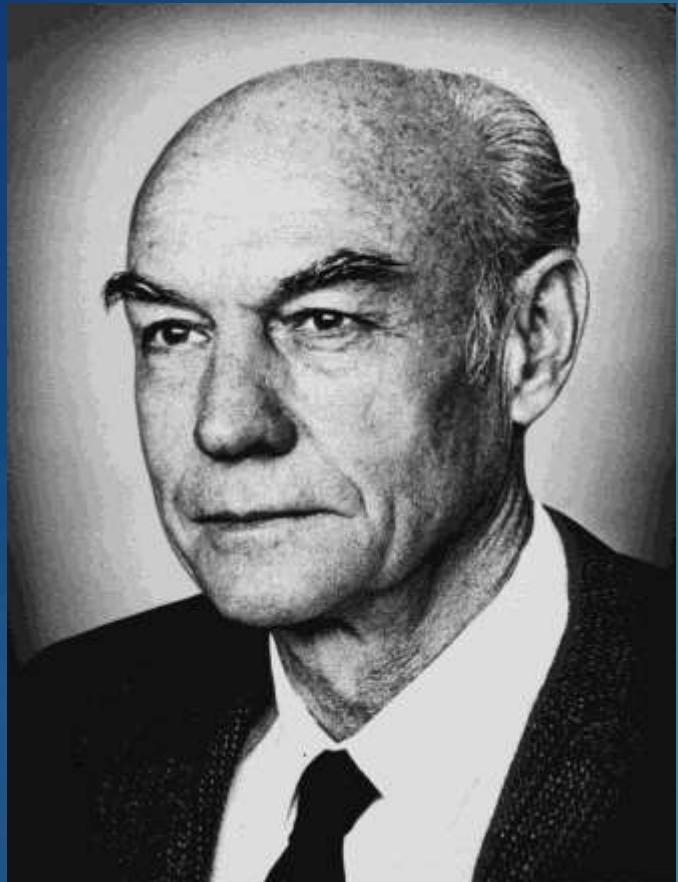
$$\begin{array}{ccc} \boxed{\text{red}} & + & \boxed{\text{yellow}} \\ & & = \\ & & \boxed{\text{yellow}} \end{array}$$

$$\begin{array}{ccc} \boxed{\text{red}} & + & \boxed{\text{cyan}} \\ & & = \\ & & \boxed{\text{purple}} \end{array}$$

$$\begin{array}{ccc} \boxed{\text{cyan}} & + & \boxed{\text{yellow}} \\ & & = \\ & & \boxed{\text{green}} \end{array}$$

emergentyzm

- ▶ **emergo (emergere)** – wyłaniam się, powstaję
- ▶ wyłonienie się jakościowo nowych własności wyższego rzędu z oddziaływań na poziomie niższym
- ▶ „całość jest czymś więcej niż sumą części”
 - ▶ nowe własności w agregatach, nieobecne w częściach składowych
 - ▶ jeśli makrocechy nie da się wyrowadzić ze znajomości mikrocech, to jest ona własnością emergentną



Richard Sperry
(1913-1994)



Tim Crane
(ur. 1962)

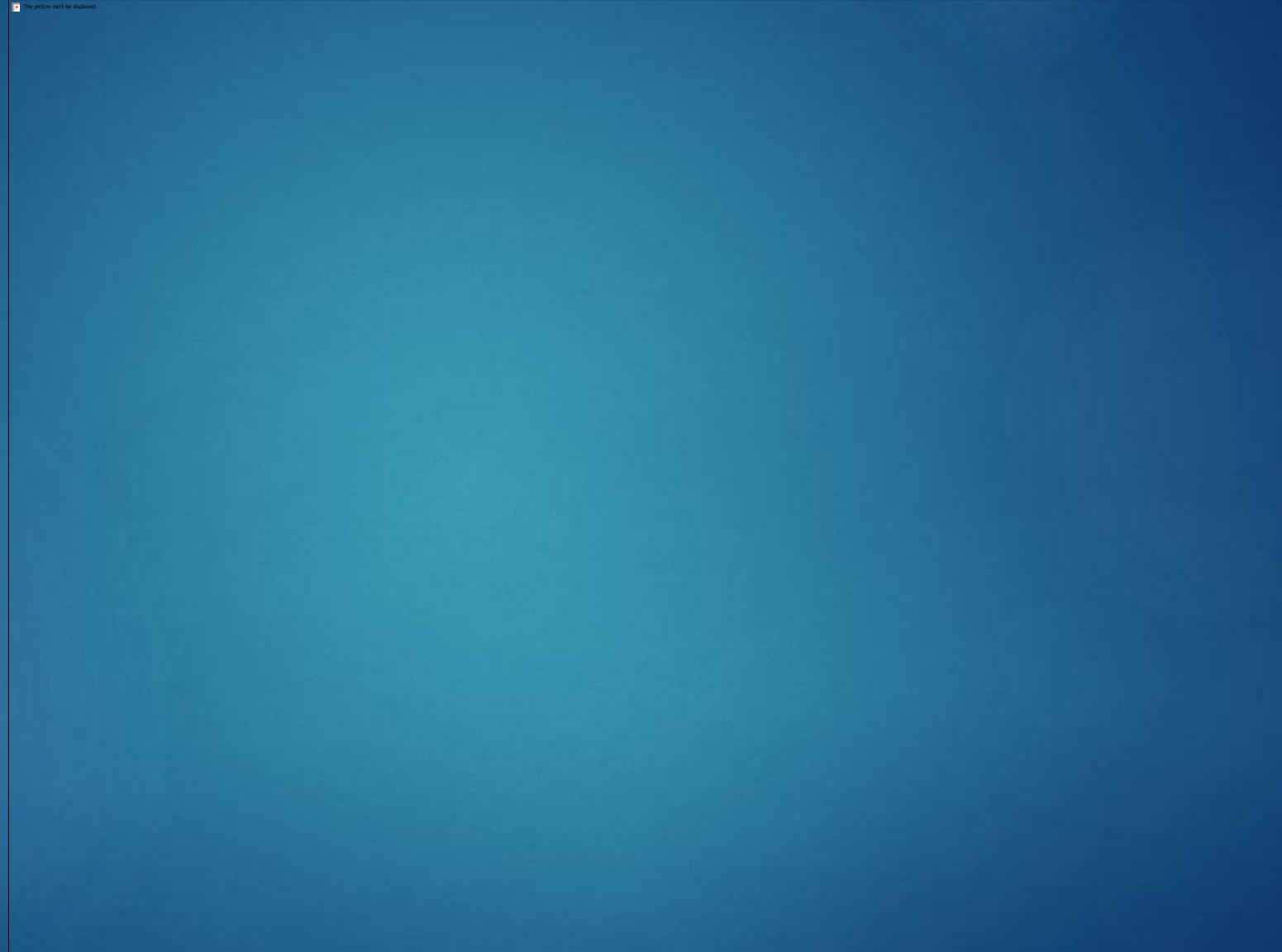
emergentyzm

- ▶ umysł/świadomość jako jakości emergentne
- ▶ funkcjonalna redukcja (rozkład złożonych funkcji na operacje elementarne) oraz emergencja są komplementarne
- ▶ wspólnie prowadzą do introspecyjnego poczucia jedności świadomego doświadczenia
- ▶ właściwości emergentne są jednak niereduksowalne ontologicznie
- ▶ właściwości emergentne są także nieprzewidywalne
- ▶ teza o przyczynowości odgórnej
- ▶ emergencja kontekstowa
 - ▶ neurofizjologia, moduły, obliczeniowość konieczne do interakcji ze środowiskiem, ale niewystarczające
 - ▶ np. emocje społeczne, nie wystarcza określenie aktywności mózgu etc.

monizm neutralny



Bertrand Russell
(1872-1970)



monizm neutralny

- ▶ alternatywa między materializmem i dualizmem jest fałszywa
 - ▶ tradycja brytyjskiego empiryzmu (Hume, Berkeley)
 - ▶ nie doświadczamy ani materii, ani substancji umysłowej
 - ▶ elementem naszego doświadczenia są perceptory (spostrzeżenia)

monizm neutralny

- ▶ alternatywa między materializmem i dualizmem jest fałszywa
- ▶ substancialna teoria jaźni
 - ▶ umysł jak naczynie, w którym znajdują się poszczególne stany mentalne
 - ▶ to kartezjańskie wyobrażenie jest błędne

monizm neutralny

- ▶ alternatywa między materializmem i dualizmem jest fałszywa
- ▶ substancialna teoria jaźni
- ▶ pierwotne elementy świata są ontologicznie neutralne
 - ▶ nie są ani fizyczne, ani psychiczne
 - ▶ to wyobrażenie jest konsekwencją dualistycznego języka
 - ▶ języka tego używamy dla wygody
 - ▶ psychologia i fizyka nie badają różnych aspektów świata, ale to samo pod różnym kątem

panpsychizm





Galen Strawson
(ur. 1952)



Philip Goff

panpsychizm

„[Możliwe jest] nadać odmienny bieg rozumowaniu, czyniąc materię strukturą złożoną z jednostek psychicznych. Nie jestem zupełnie pewny, że pogląd taki byłby fałszywy.”

Bertrand Russell, *Zarys filozofii*

panpsychizm

- ▶ przyjęcie fizykalizmu pociąga za sobą panpsychizm
 - ▶ materialiści uznają, że każda cząstka Wszechświata jest materialna/fizyczna
 - ▶ przyjmują oni również istnienie stanów psychicznych/świadomości
 - ▶ zatem odpowiednie ułożenie materii skutkuje umysłowością
 - ▶ ale materia składa się z tych samych elementów (leptony, kwarki, bozony...)
 - ▶ zatem każda materia, przy odpowiedniej konfiguracji, ma potencjał stania się umysłową
 - ▶ wniosek: każda cząstka Wszechświata ma (częściowo) charakter umysłowy

wstęp do filozofii umysłu i sztucznej inteligencji

30.11.2023

WYKŁAD 8

panpsychizm

- ▶ przyjęcie fizykalizmu pociąga za sobą panpsychizm
 - ▶ materialiści uznają, że każda cząstka Wszechświata jest materialna/fizyczna
 - ▶ przyjmują oni również istnienie stanów psychicznych/świadomości
 - ▶ zatem odpowiednie ułożenie materii skutkuje umysłowością
 - ▶ ale materia składa się z tych samych elementów (leptony, kwarki, bozony...)
 - ▶ zatem każda materia, przy odpowiedniej konfiguracji, ma potencjał stania się umysłową
 - ▶ wniosek: każda cząstka Wszechświata ma (częściowo) charakter umysłowy

panpsychizm

- ▶ przyjęcie fizykalizmu pociąga za sobą panpsychizm
- ▶ teoria podwójnego aspektu
 - ▶ materialność i umysłowość są dwoma aspektami tej samej rzeczywistości
 - ▶ materialność i umysłowość są niereduksywne
 - ▶ nigdy nie postrzegamy materii, a jedynie nasze wrażenia (empiryzm)

panpsychizm

- ▶ przyjęcie fizykalizmu pociąga za sobą panpsychizm
- ▶ teoria podwójnego aspektu
- ▶ wnioskowanie abdukcyjne (do najlepszego wyjaśnienia)
 - ▶ fizyka nie mówi nic o wewnętrznych własnościach materii
 - ▶ istnienie umysłowości/świadomości jest pewną niezależną daną
 - ▶ panpsychizm dostarcza najlepszego spójnego wyjaśnienia

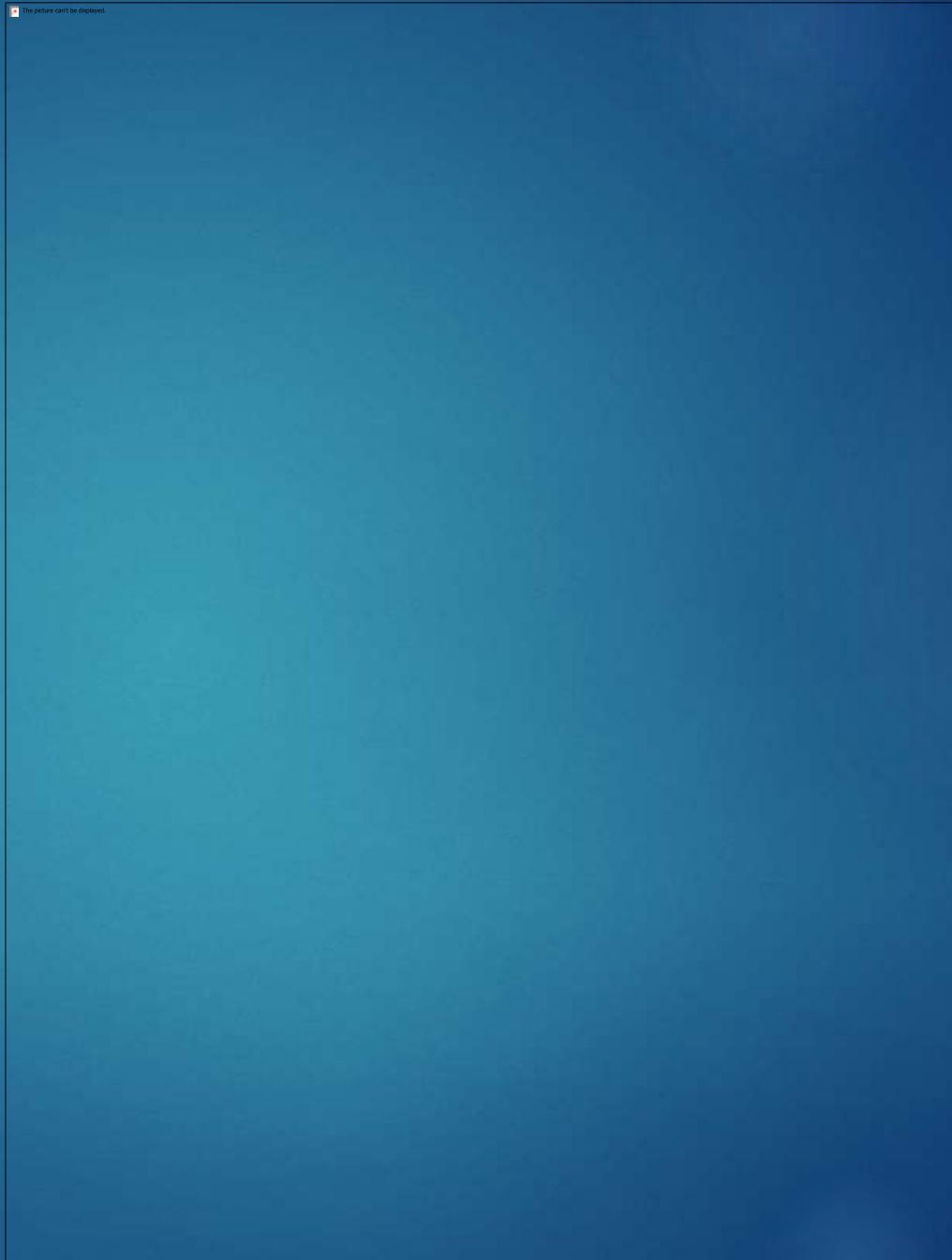
panpsychizm

- ▶ przyjęcie fizykalizmu pociąga za sobą panpsychizm
- ▶ teoria podwójnego aspektu
- ▶ wnioskowanie abdukcyjne (do najlepszego wyjaśnienia)
- ▶ paneksperientyzm
 - ▶ *qualia* jako fundamentalna własność psychiczna
 - ▶ doświadczenia jakościowe mogą być odczuwane tylko na pewnym poziomie złożenia materii
 - ▶ musi istnieć podmiot
 - ▶ paneksperientyzm ≠ pansubiektywizm

naturalizm biologiczny



John Searle
(ur. 1932)



naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
 - ▶ dualizm prowadzi do absurdu, jego motywacje mają charakter irracjonalny (religijny, spirytualistyczny)
 - ▶ materializm ignoruje istotne własności umysłu
 - ▶ materializm wciąż używa dualistycznych kategorii (umysł, materia, mentalne, fizyczne) → powinniśmy je odrzucić bądź zredefiniować
 - ▶ naturalizm metodologiczny → odrzucenie „dualizmu pojęciowego”

naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
- ▶ przekonanie o niedostępności poznawczej umysłu (subiektywnego) dla nauki (obiektywnej) opiera się na błędzie kategorialnym
 - ▶ dwa rozumienia subiektywności i obiektywności: epistemologiczne i ontologiczne

naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
- ▶ przekonanie o niedostępności poznawczej umysłu (subiektywnego) dla nauki (obiektywnej) opiera się na błędzie kategorialnym
 - ▶ dwa rozumienia subiektywności i obiektywności: **epistemologiczne** i ontologiczne

Pies Johna Searle'a ma na imię Tarski. → epistemologicznie obiektywne

Pies Johna Searle'a jest bardzo piękny. → epistemologicznie subiektywne

naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
- ▶ przekonanie o niedostępności poznawczej umysłu (subiektywnego) dla nauki (obiektywnej) opiera się na błędzie kategorialnym
 - ▶ dwa rozumienia subiektywności i obiektywności: epistemologiczne i ontologiczne

John Searle i pies Tarski istnieją w sposób **obiektywny**.

Śwędzenie nosa Johna Searle'a i ból łapy psa Tarskiego istnieją w sposób **subiektywny**.

naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
- ▶ przekonanie o niedostępności poznawczej umysłu (subiektywnego) dla nauki (obiektywnej) opiera się na błędzie kategorialnym
 - ▶ dwa rozumienia subiektywności i obiektywności: epistemologiczne i ontologiczne
 - ▶ stany umysłowe są ontologicznie subiektywne, ale epistemologicznie obiektywne
 - ▶ nauka zajmuje się badaniem tego, co epistemologicznie obiektywne

naturalizm biologiczny

- ▶ umysł/świadomość jako część biologicznego wyposażenia
- ▶ odrzucenie dualizmu i materializmu
- ▶ przekonanie o niedostępności poznawczej umysłu (subiektywnego) dla nauki (obiektywnej) opiera się na błędzie kategorialnym
- ▶ w celu stworzenia adekwatnej teorii umysłu należy dopuścić wielopoziomowość opisów



wstęp do filozofii umysłu i sztucznej inteligencji

7.12.2023

WYKŁAD 9

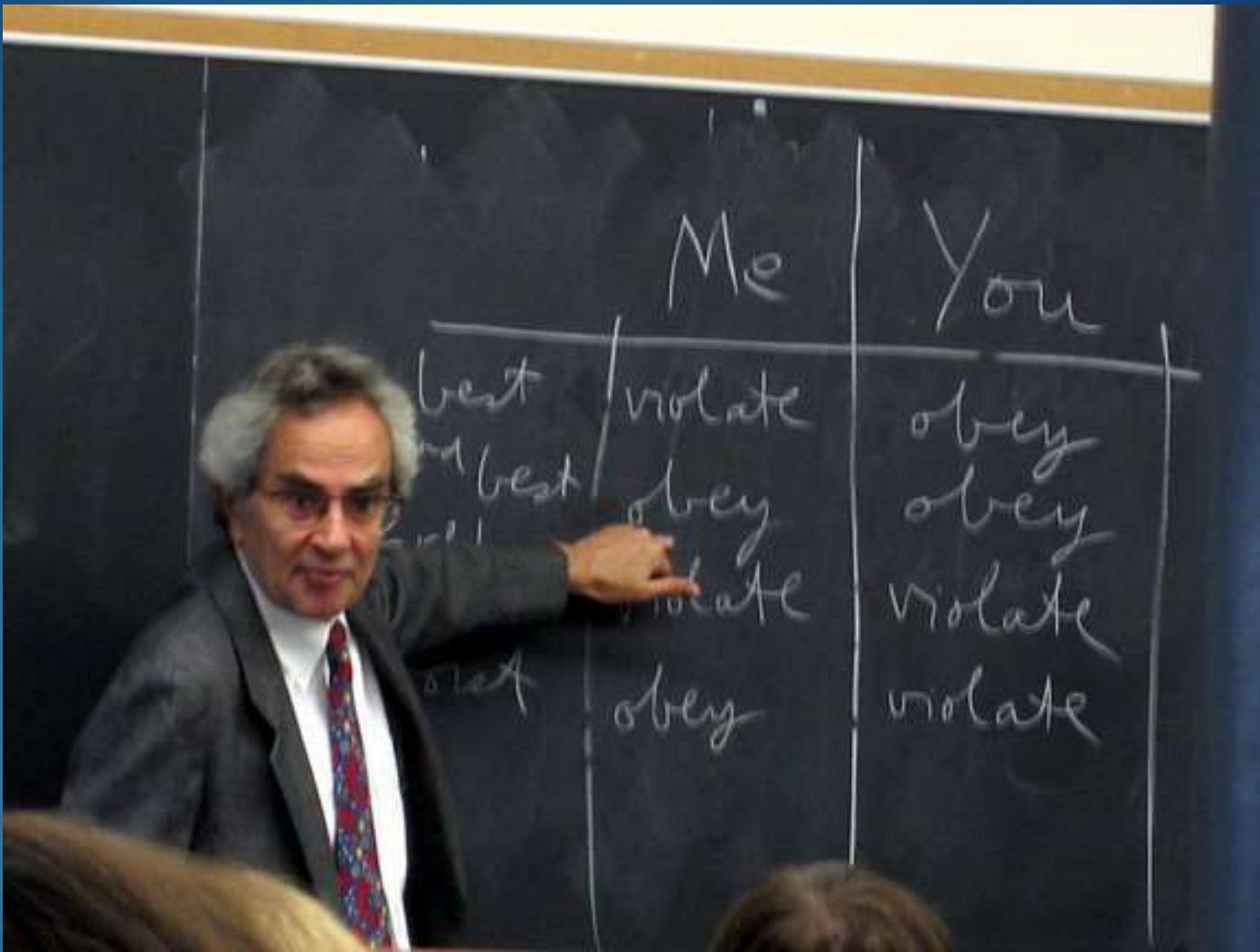
świadomość





To świadomość czyni ze stosunku umysłu do ciała coś, z czym naprawdę trudno się uporać. [...] Gdyby nie istniała świadomość, problem stosunku umysłu do ciała byłby mniej interesujący. Ze świadomością – wydaje się beznadziejny.

Thomas Nagel



Thomas Nagel
(ur. 1937)





To świadomość czyni ze stosunku umysłu do ciała coś, z czym naprawdę trudno się uporać. [...] Gdyby nie istniała świadomość, problem stosunku umysłu do ciała byłby mniej interesujący. Ze świadomością – wydaje się beznadziejny.

Thomas Nagel, *Jak to jest być nietoperzem?*, 1974

świadomość – wyliczanka pojęciowa

- ▶ świadomość jako samoświadomość
- ▶ świadomość jako przytomność
- ▶ świadomość jako „bycie na jawie”
- ▶ świadomość jako uświadamianie sobie czegoś
- ▶ świadomość jako zdolność do introspekcji
- ▶ świadomość jako jaźń
- ▶ świadomość jako subiektywna jakość przeżyć
 - ▶ doznania/przeżycia, qualia, fenomenologia
 - ▶ jak to jest być kimś

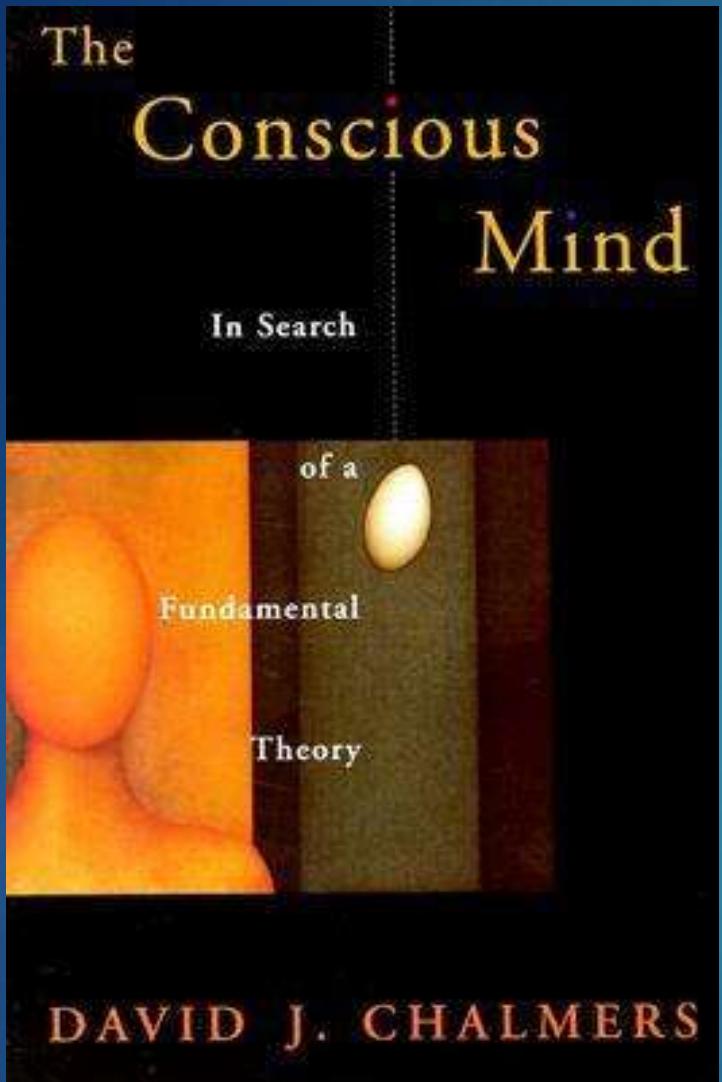
świadomość – wyliczanka pojęciowa

Kiedy postrzegamy, myślimy i działamy, powstaje wir związków przyczynowych i przetwarzania informacji, ale to przetwarzanie nie zachodzi zwykle w ukryciu. Istnieje także aspekt wewnętrzny, jest coś takiego jak poczucie bycia podmiotem poznawczym. **Ten wewnętrzny aspekt to przeżycie świadome.** Przeżycia świadome obejmują rozmaite zjawiska od żywych doznań barw po przeżycia najsłabiejszych woni w tle; od ostrych bólów po ulotne przeżycie myśli na końcu języka; od zwykłych dźwięków i zapachów po wszechobejmującą wzniósłość przeżycia muzycznego; od banalności dokuczliwego swędzenia po głębie lęków egzystencjalnych; od konkretności smaku mięty pieprzowej po ogólność własnego przeżycia tożsamości. Wszystkie te przeżycia mają różne przeżywane jakości. Wszystkie stanowią istotne elementy wewnętrznego życia umysłu.

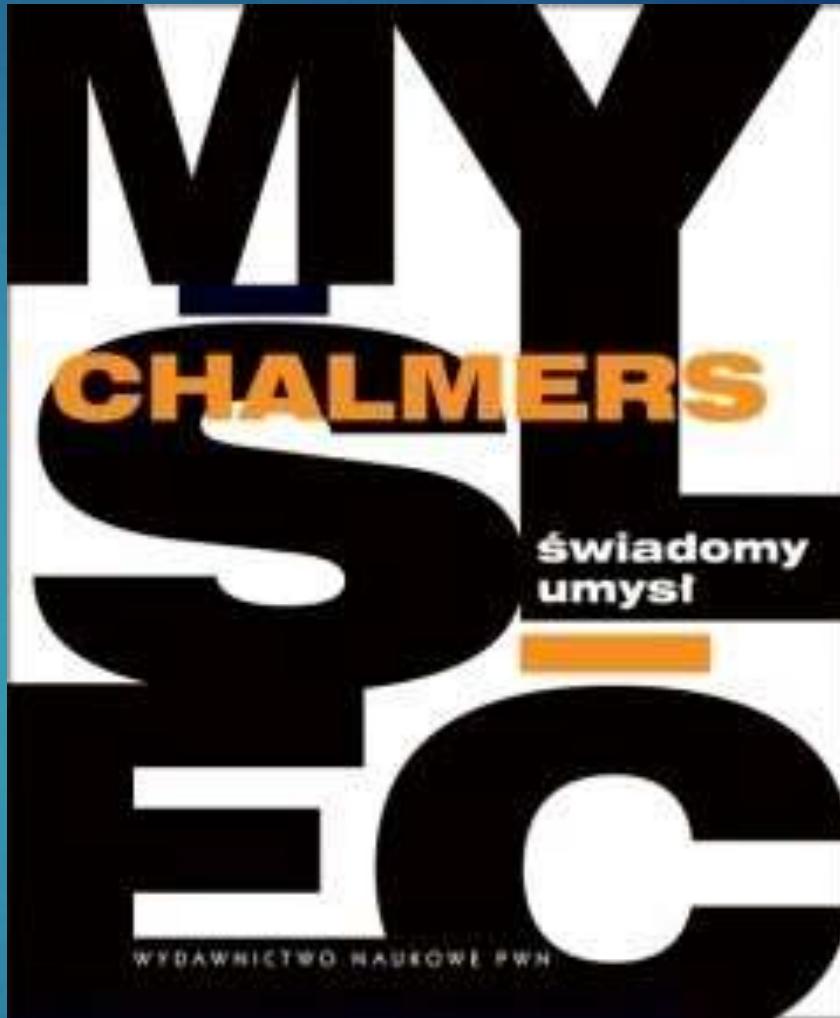
David Chalmers, Świadomy umysł (1996)



David Chalmers
(ur. 1966)



1996



2010

świadomość – wyliczanka pojęciowa

- ▶ katalog przeżyć świadomych według Chalmersa
 - ▶ przeżycia wzrokowe
 - ▶ przeżycia słuchowe
 - ▶ przeżycia dotykowe
 - ▶ przeżycia węchowe
 - ▶ przeżycia smakowe
 - ▶ przeżycia ciepła i zimna
 - ▶ ból
 - ▶ inne doznania cielesne
 - ▶ obrazy umysłowe
 - ▶ myśl świadoma
 - ▶ emocje
 - ▶ poczucie samego siebie

świadomość – wyliczanka pojęciowa

- ▶ katalog przeżyć świadomych według Chalmersa
- ▶ fenomenalna i psychologiczna koncepcja umysłu
 - ▶ koncepcja fenomenalna kładzie nacisk na jakościowy aspekt umysłu
 - ▶ w koncepcji psychologicznej umysł to eksplanacyjna podstawa zachowań
 - ▶ fenomenalna: jak się umysł odczuwa, psychologiczna: co umysł robi

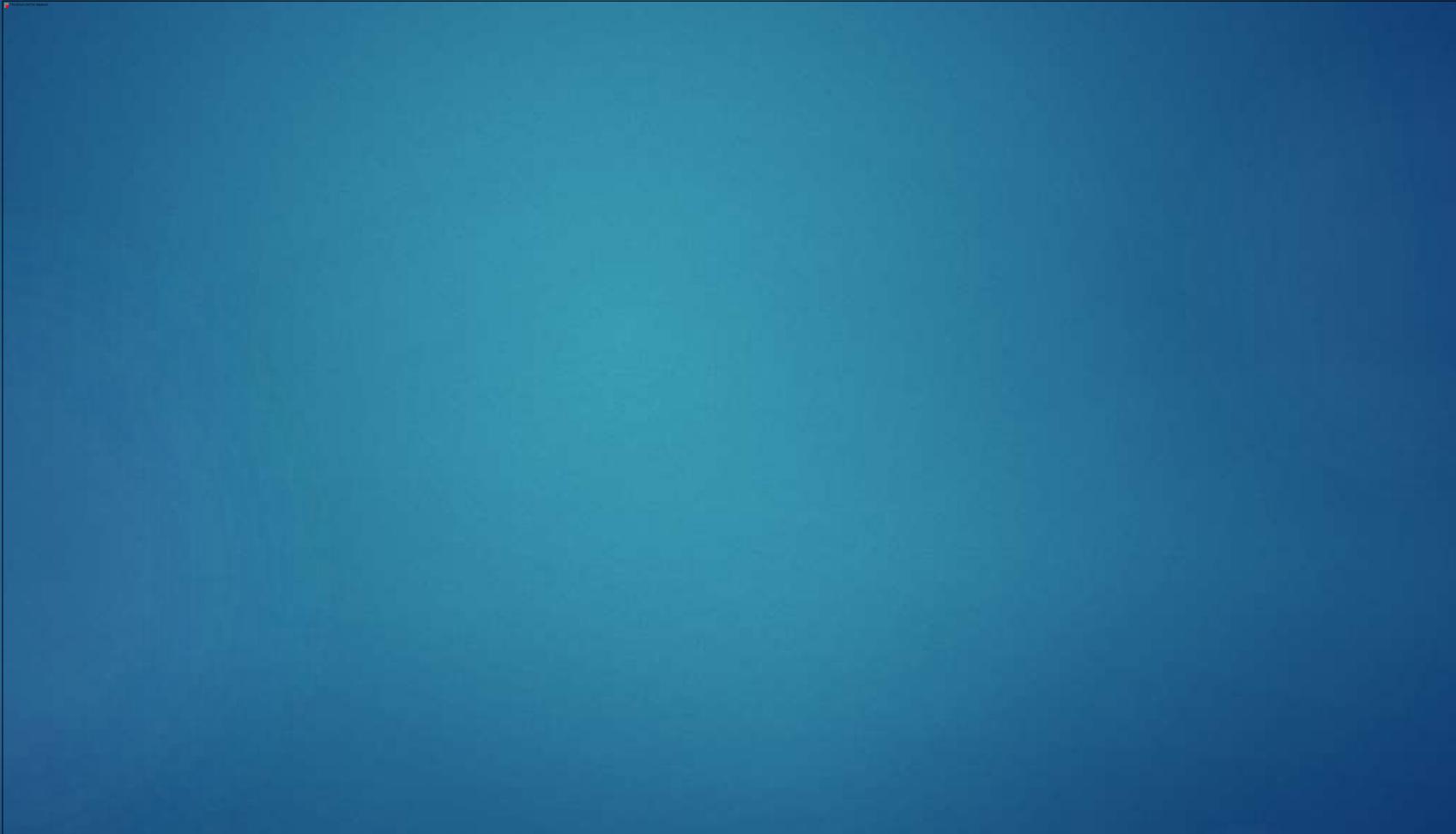
łatwe problemy i Trudny Problem Świadomości

- ▶ łatwe problemy świadomości
 - ▶ dotyczyć będą psychologicznej koncepcji umysłu
 - ▶ np. „Jak układ fizyczny mógłby być tego rodzaju rzeczą, która byłaby w stanie uczyć się lub pamiętać?”
 - ▶ pytania te dotyczą ostatecznie własności funkcjonalnych i/lub przyczynowych
 - ▶ łatwe problemy wcale nie są łatwe → ale **co do zasad** wiemy, jak się za nie zabrać (nie stwarzają poważnych problemów metafizycznych)
 - ▶ są to problemy, które można próbować rozwiązywać np. na gruncie kognitywistyki, psychologii poznawczej, neuronauk...

łatwe problemy i Trudny Problem Świadomości

- ▶ łatwe problemy świadomości
- ▶ Trudny Problem Świadomości
 - ▶ dotyczy fenomenalnej koncepcji umysłu
 - ▶ generalne pytanie: dlaczego właściwościom psychologicznym towarzyszą właściwości fenomenalne?
 - ▶ w tym sensie TPŚ to problem umysł-umysł
 - ▶ np. „Dlaczego wszystkim bodźcom i reakcjom związanym z bólem towarzyszy przeżycie bólu?”
 - ▶ łatwe problemy to rozmaite (często niełatwne!) zagadki, natomiast TPŚ jest tajemnicą

filozoficzne zombie



filozoficzne zombie



Chalmers



Zombie-Chalmers

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć



pod kątem fizycznym, funkcjonalnym, psychologicznym...

(Nie jest być może niczym zaskakującym, że [filozoficzne zombie] nie cieszyły się popularnością w Hollywood – byłyby oczywiste problemy z ich przedstawieniem.)

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć



nie ma żadnego *jak to jest być zombie*

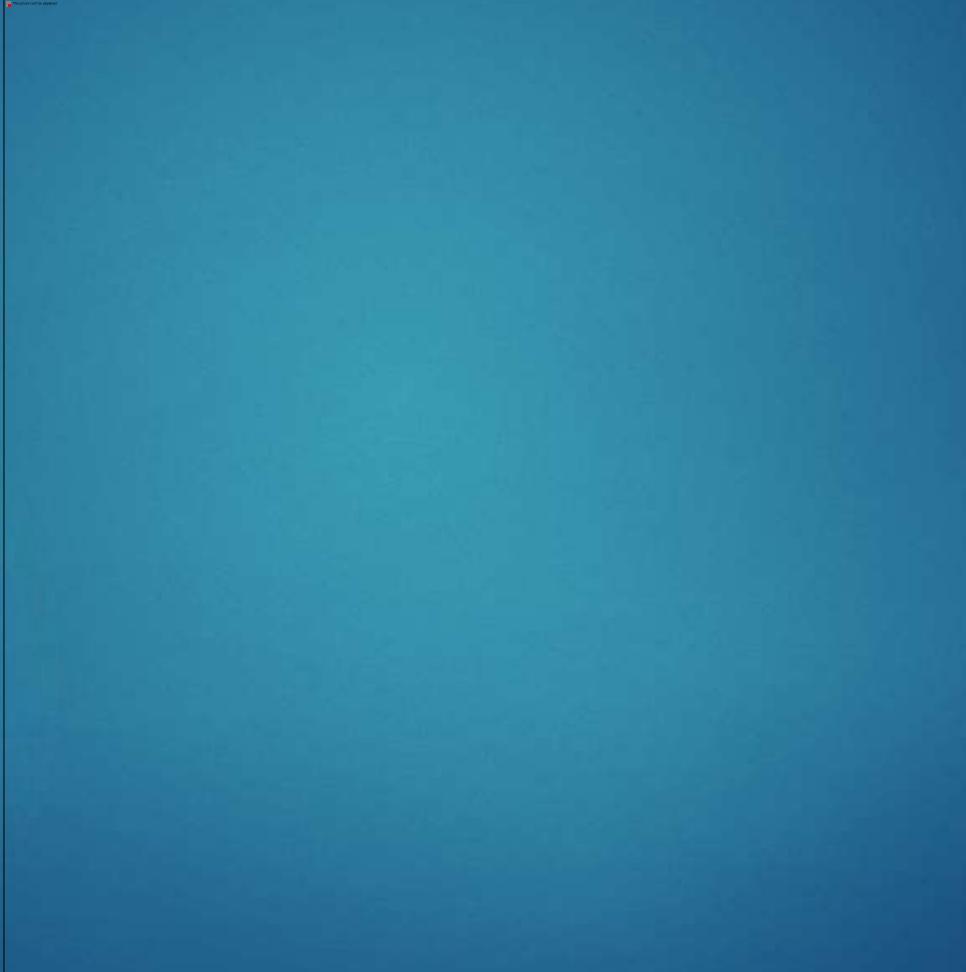
filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć



pytanie poboczne: czy zombie miałyby (ludzką) podmiotowość?

filozoficzne zombie



filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomości przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
 - ▶ nie jest (zapewne) możliwy empirycznie
 - ▶ ale nie pytamy o prawdopodobieństwo jego zaistnienia, ale spójność pojęciową opisu
 - ▶ skoro opis zombie nie zawiera żadnej sprzeczności i zombie jest pojmowalny, to jest tym samym logicznie możliwy

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu

P₁. Jeśli materializm jest prawdziwy, to przeżycia są tożsame ze stanami mózgu

P₂. Jeśli zombie są logicznie możliwe, to logicznie możliwe jest, by stany mózgu istniały bez przeżyć

P₃. Zatem jeśli zombie są logicznie możliwe, materializm jest fałszywy

P₄. Zombie są logicznie możliwe

W. Materializm jest fałszywy

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu

P₁. Jeśli fizykalizm jest prawdziwy, to niemożliwe jest, by dwa identyczne pod względem wszystkich fizycznych faktów światy różniły się

P₂. Jeśli zombie są logicznie możliwe, to logicznie możliwe jest, by dwa identyczne pod względem wszystkich fizycznych faktów światy różniły się

P₃. Zatem jeśli zombie są logicznie możliwe, fizykalizm jest fałszywy

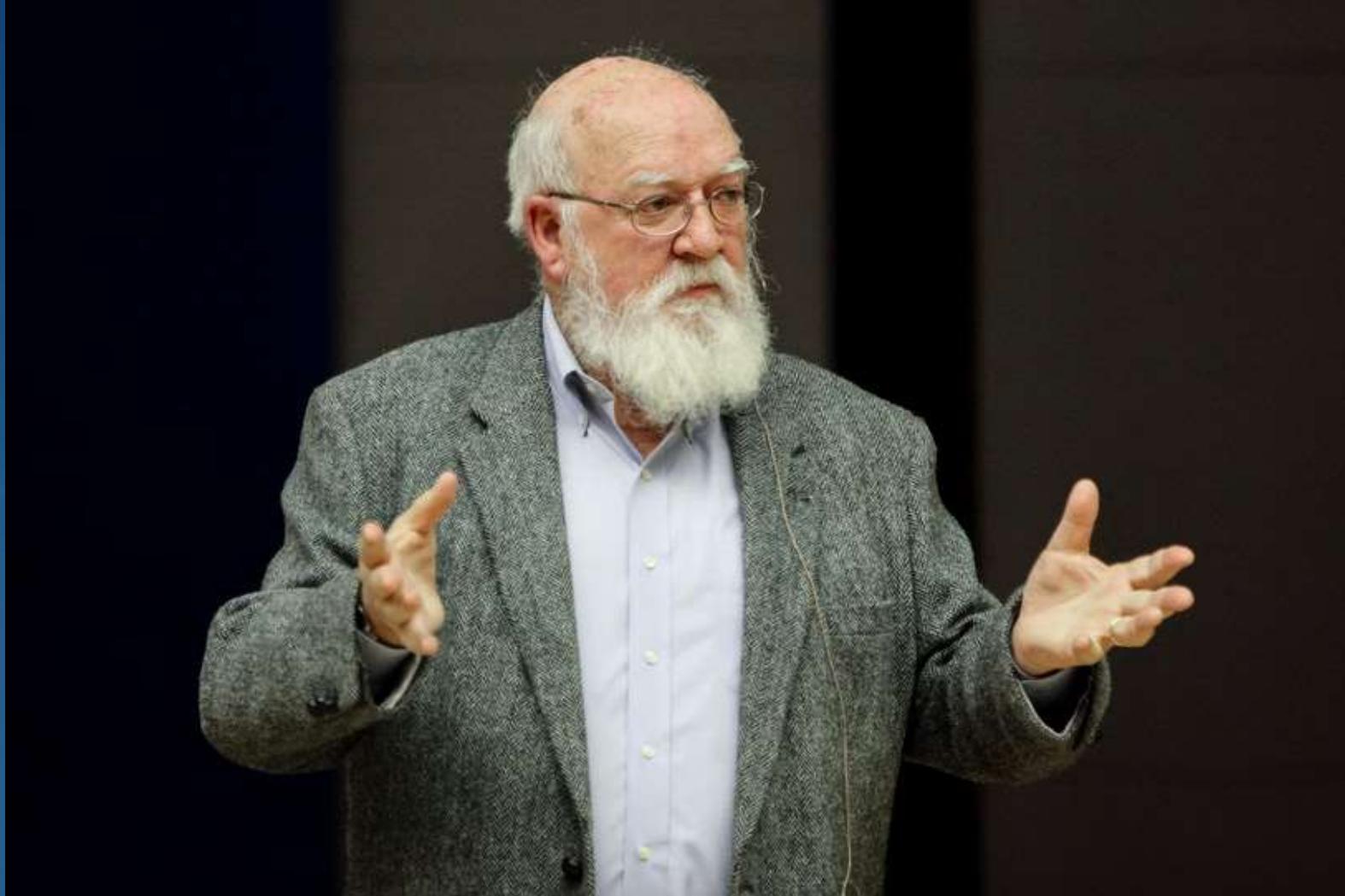
P₄. Zombie są logicznie możliwe

W. Fizykalizm jest fałszywy

filozoficzne zombie

Argument z możliwości zombie okrężną drogą dowodzi tego, co właściwie wydaje się oczywiste. Opisanie jakiejś osoby pod kątem natury fizycznej oznacza nadanie jej pewnych obiektywnych i ilościowych cech. Aby natomiast opisać kogoś w wymiarze doświadczeń świadomych, musimy przypisać mu cechy subiektywne i jakościowe. To, że system fizyczny ma pierwszy rodzaj cech, nie oznacza, że ma też drugi. **Oto właśnie rdzeń problemu świadomości.** Współczesny materializm nie jest jego rozwiązaniem, ale upartym zaprzeczaniem, że jakakolwiek trudność w ogóle tu zachodzi.

Philip Goff, Błąd Galileusza, PWN 2022



Daniel Dennett
(ur. 1942)



Wymyślanie przykładów, które „udowadniają” jakieś kwestie pojęciowe, to zabawa dla dzieci. Założmy, że krowa urodziła coś, co na poziomie atomów było nieodróżnialne od rekina. Czy byłby to rekin? (...) Lub założmy, że zły demon sprawił, iż woda przechodzi w stały stan skupienia w temperaturze pokojowej w wyniku uśmiechnięcia się do niej; czy ta demoniczna woda byłaby lodem? (...) Uśmiechające się demony, kroworekiny, zombie i ludzie z bagien są, jak uważają niektórzy filozofowie, logicznie możliwe (...) i filozofowie ci uważają, że jest to ważne.

Daniel Dennett, *Dźwignie wyobraźni
i inne narzędzia do myślenia*, CCP 2015

Pozwól mi od razu przyznać, że moje Poczucie Sensowności Zombie jest tak samo silne jak u każdego. Gdy patrzę na to z przymrużonymi oczami i pod odpowiednim kątem, to oczywiście wydaje mi się, że świadomość musi być czymś więcej niż te wszystkie rzeczy, które robi dla nas i nam, jakimś rodzajem szczególnego, prywatnego blasku lub jestem-tutaj-owości (...). Ale nauczyłem się nie zwracać uwagi na to przeczucie. Myślę, że jest ono oczywistym błędem, porażką wyobraźni, a nie wglądem w konieczność.

Daniel Dennett, Dźwignie wyobraźni
i inne narzędzia do myślenia, CCP 2015

wstęp do filozofii umysłu i sztucznej inteligencji

14.12.2023

WYKŁAD 10

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu

P₁. Jeśli materializm jest prawdziwy, to przeżycia są tożsame ze stanami mózgu

P₂. Jeśli zombie są logicznie możliwe, to logicznie możliwe jest, by stany mózgu istniały bez przeżyć

P₃. Zatem jeśli zombie są logicznie możliwe, materializm jest fałszywy

P₄. Zombie są logicznie możliwe

W. Materializm jest fałszywy

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu

P₁. Jeśli fizykalizm jest prawdziwy, to niemożliwe jest, by dwa identyczne pod względem wszystkich fizycznych faktów światy różniły się

P₂. Jeśli zombie są logicznie możliwe, to logicznie możliwe jest, by dwa identyczne pod względem wszystkich fizycznych faktów światy różniły się

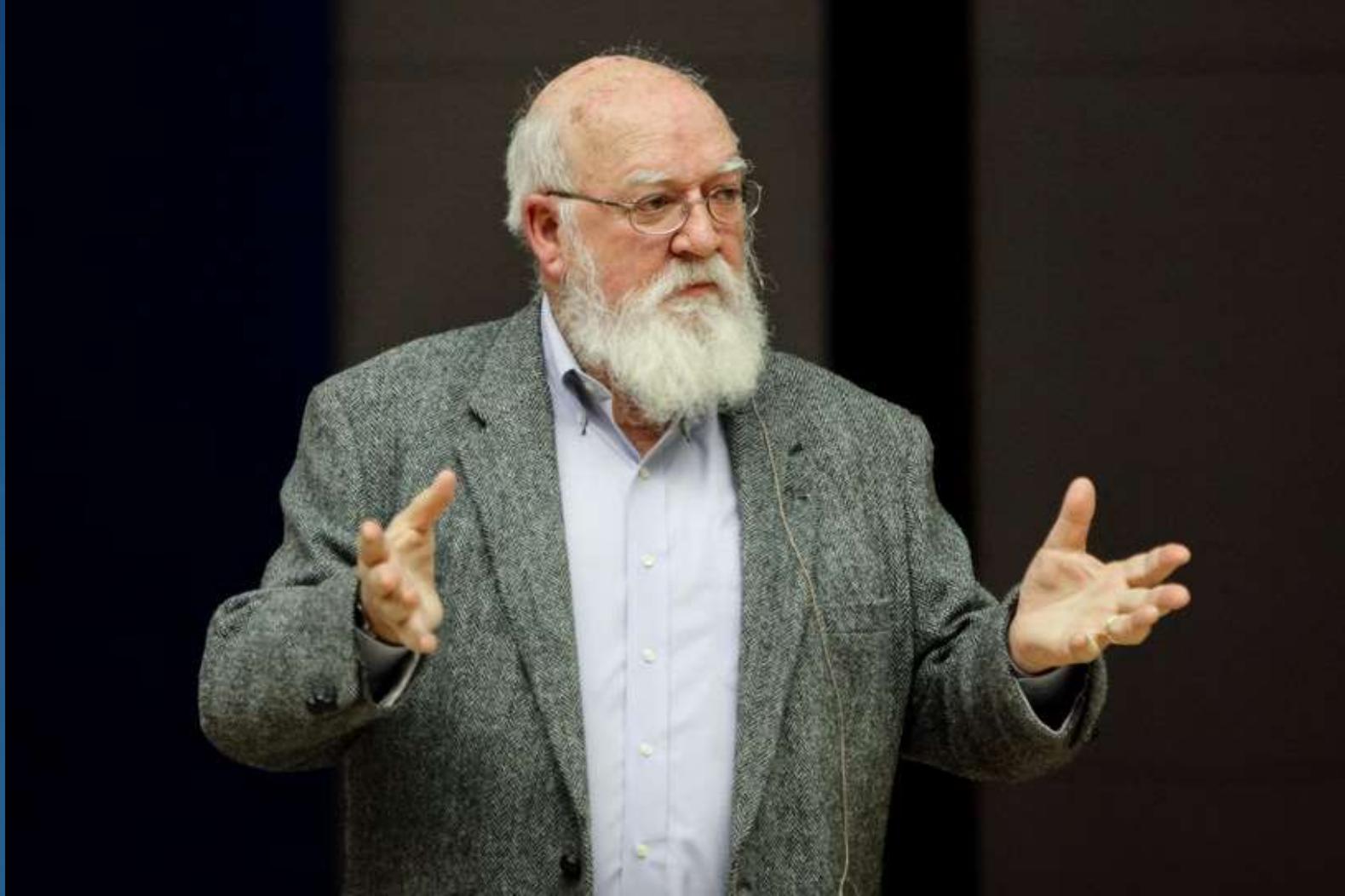
P₃. Zatem jeśli zombie są logicznie możliwe, fizykalizm jest fałszywy

P₄. Zombie są logicznie możliwe

W. Fizykalizm jest fałszywy

filozoficzne zombie

- ▶ zombie – ktoś/coś identyczne ze mną, ale pozbawione świadomych przeżyć
- ▶ zombie jest logicznie możliwy, bo jego opis jest niesprzeczny
- ▶ logiczna możliwość zombie prowadzi do odrzucenia materializmu/fizykalizmu
- ▶ potrzebujemy zatem materializmu+/fizykalizmu+
 - ▶ oprócz fundamentalnych praw fizyki potrzebujemy dodatkowych praw fundamentalnych



Daniel Dennett
(ur. 1942)

filozoficzne zombie

Jeśli wierzymy, że świadomość dzieli wszechświat na pół, na rzeczy, które ją mają, i na te, którym jej brak, oraz sądzimy, że jest to podstawowe rozróżnienie metafizyczne, wówczas żądanie fundamentalnych praw [...] ma jakiś sens. My, naturaliści, sądzimy jednak, że takie przebóstwianie świadomości samo w sobie jest podejrzane, wspiera się tylko na tradycji i na niczym więcej. Zauważmy, że nikt dzisiaj nie woła o fundamentalne prawa teorii kangurów, które pokazywałyby, iż pseudokangury są fizycznie, logicznie, metafizycznie niemożliwe. Kangury są cudowne, lecz nie aż tak cudowne.

Daniel Dennett, Słodkie sny, Prószyński i S-ka 2007



Gottfried Wilhelm Leibniz

(1646-1716)



zombie a zimbo

- ▶ pojmowanie a wyobrażanie sobie
 - ▶ czy jesteśmy w stanie pojąć tysiącbok?
 - ▶ czy na pewno wiemy, co to oznacza pojąć możliwość filozoficznego zombie?

zombie a zimbo

- ▶ pojmowanie a wyobrażanie sobie
- ▶ zimbo
 - ▶ szczególny podgatunek zombie
 - ▶ ma stany informacyjne wyższego rzędu
 - ▶ potrafi monitorować swoje stany mentalne
 - ▶ ...oraz monitorować ten monitoring
 - ▶ wyposażony jest zatem w *nieświadomą samoreprezentację rekursywną* („jeśli to ma jakiś sens”)
 - ▶ innymi słowy: zimbo to *refleksyjny zombie* (ale wciąż nieświadomy!)

zombie a zimbo

- ▶ pojmowanie a wyobrażanie sobie
- ▶ zimbo
- ▶ jak odróżnić zimbo od zombie?
 - ▶ jeśli zombie pozbawiony byłby takiej samorefleksji, to jak zdawałby raporty ze swoich (nieświadomych) stanów mentalnych?

zombie a zimbo

- ▶ pojmowanie a wyobrażanie sobie
- ▶ zimbo
- ▶ jak odróżnić zimbo od zombie?
- ▶ jak odróżnić zimbo od świadomego człowieka?
 - ▶ jeśli świadomość wiąże się ze zdolnością do samorefleksji, to zimbo jest świadomym
 - ▶ jeśli nie wiąże się, to nie ma możliwości zaproponowania rozstrzygającego testu
 - ▶ jeśli różnica między zimbo/zombie a człowiekiem jest empirycznie niewykrywalna, to sam opis jest niespójny
 - ▶ co więcej, sama idea zombie nie jest wcale pojmowalna



Jeśli nie podejmiesz trudu wyobrażenia sobie, ze szczególnymi, jak nieodróżnialny byłby „normalny” [człowiek] od zimbo [...], to nie próbowałeś tak naprawdę pojąć zombie filozofów. Jesteś jak Leibniz, poddając się i nawet nie próbując na poważnie. [...] Dlaczego miałoby cię obchodzić, czy jesteś lub stałeś się zimbo? W gruncie rzeczy nie mógłbyś tego wiedzieć.

Daniel Dennett, Dźwignie wyobraźni
i inne narzędzia do myślenia, CCP 2015

argument z wiedzy





It's called
"Mary in the Black
and White Room."

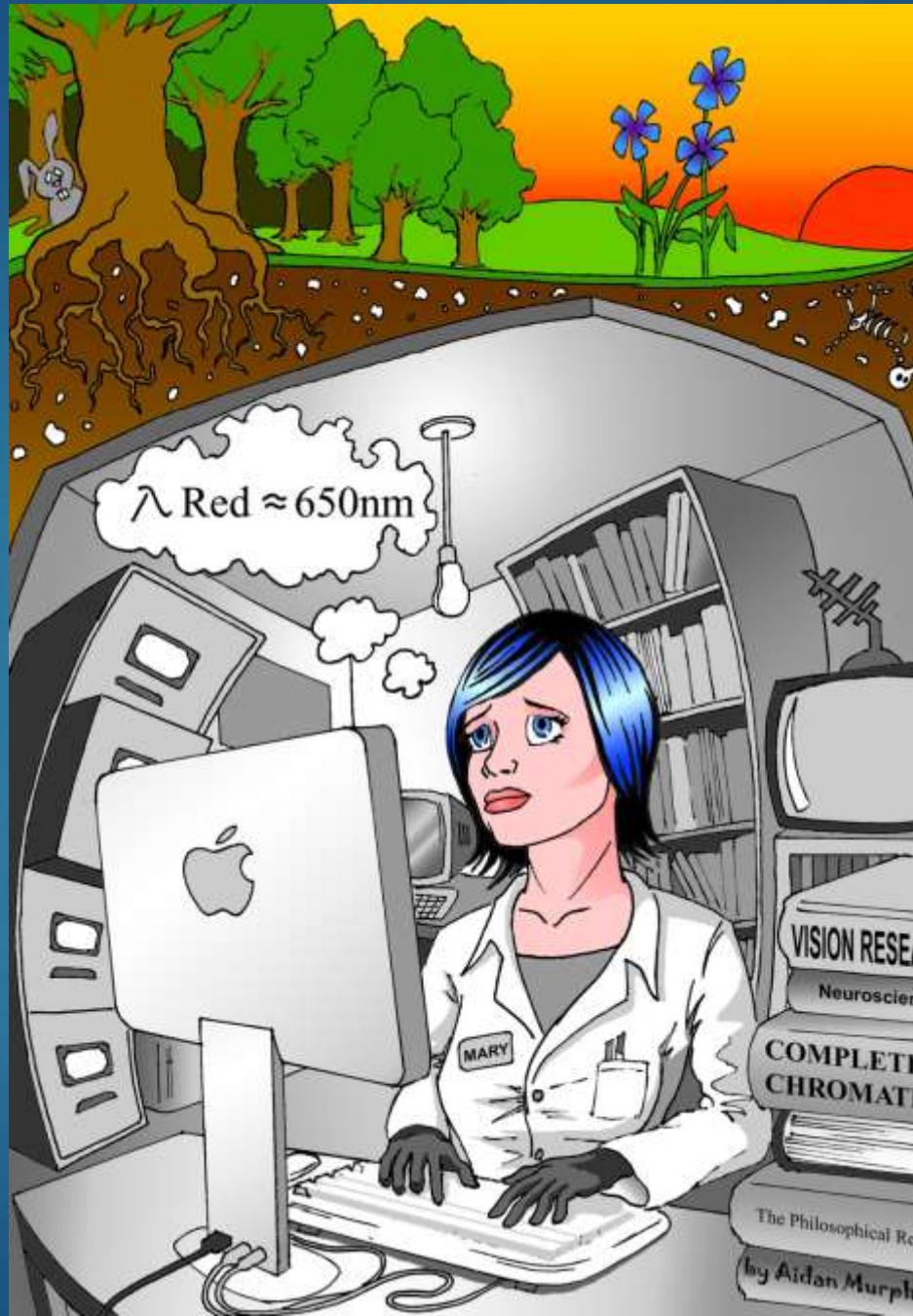
Frank Jackson
(ur. 1943)

FRANK JACKSON

*School of Social Sciences
Central University, Australia*

Department of Philosophy HK
Management Science Ltd.





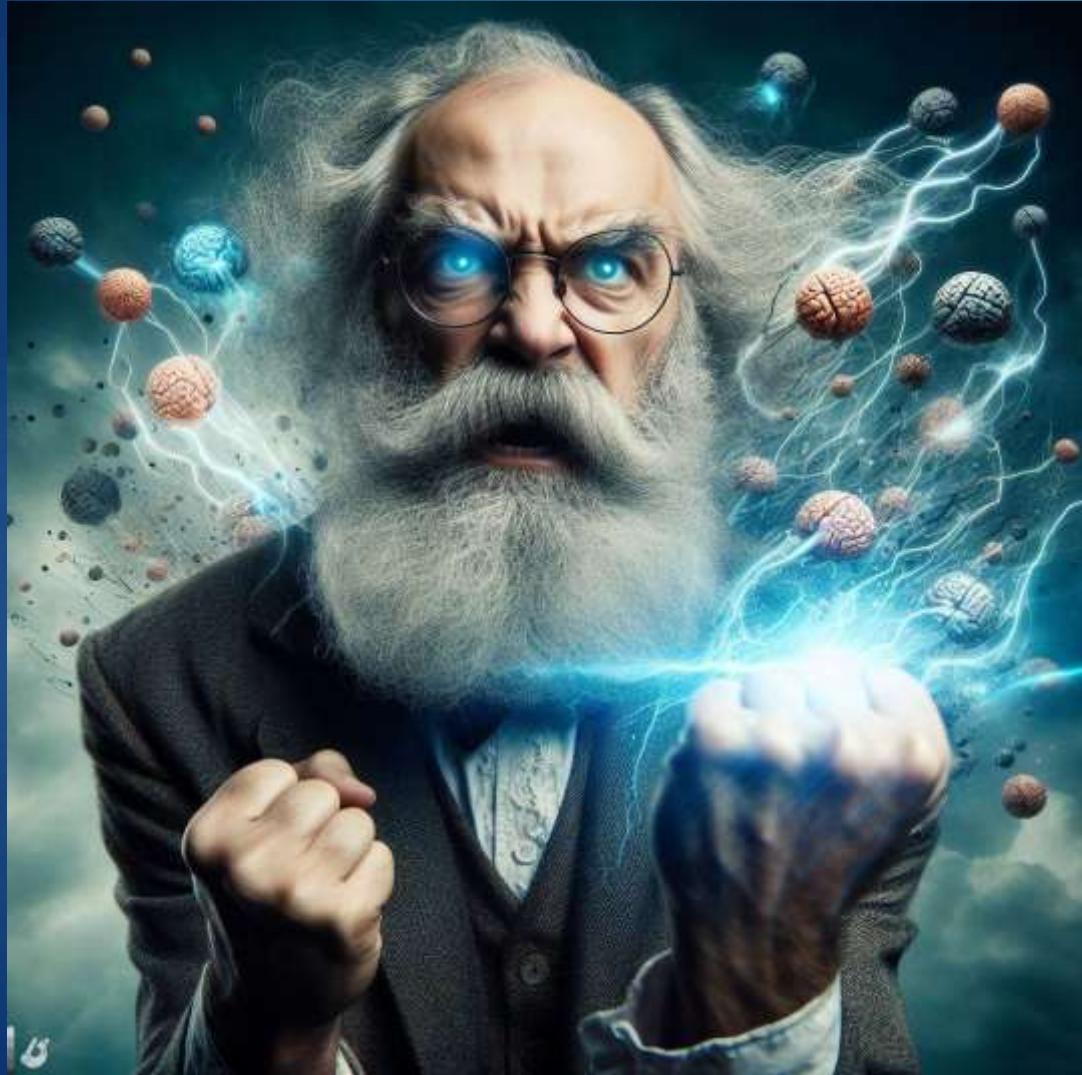
argument z wiedzy

- ▶ nawet kompletna wiedza neuronaukowa nie dostarczy nam kompletnej wiedzy o świadomości
 - ▶ Mary wie wszystko, co ma związek z neurofizjologią, a jednak dowiaduje się czegoś nowego, widząc kolor po raz pierwszy

argument z wiedzy

- ▶ nawet kompletna wiedza neuronaukowa nie dostarczy nam kompletnej wiedzy o świadomości
- ▶ a zatem istnieje luka eksplanacyjna
- ▶ tym, czego nie wyjaśnia fizykalizm, są *qualia*
- ▶ mają one charakter epifenomenalny
- ▶ ale zarazem są jakąś fundamentalną własnością świata

Dennett vs. *qualia*



Dennett vs. *qualia*

- ▶ dlaczego zdaniem Dennetta argument Jacksona jest błędny?
 - ▶ „czarno-biały pokój” → czy Mary ma rękawiczki? nie może potrzeć oczu? spojrzeć w lustro? nie może śnić kolorów?
 - ▶ „wszystkie możliwe fizyczne informacje związane z postrzeganiem kolorów” → ile to jest „wszystkie”? czy dotyczy to wszystkich mózgów, włącznie z jej? jeśli wszystkie, to jak to możliwe bez eksperymentów z kolorami?
 - ▶ „wydaje się oczywiste, że dowie się czegoś nowego o świecie” → dlaczego to oczywiste? czy tylko przy naszym stanie poznaowczym?
 - ▶ hipoteza niebieskiego banana

Dennett vs. *qualia*

- To nie było proste. W rzeczywistości wywnioskowała to za pomocą dowodu, który liczył 4765 kroków (dla barwy czerwonej – kiedy wydedukowała, jak będzie dla niej wyglądać czerwień, szybko uporała się z zielonym, przy użyciu lematu w 300 krokach, a reszta barw [...] okazała się stosunkowo banalnymi rozszerzeniami tych dowodów.
- [...]
- Zastanów się. To oczywiste! Nie można wydedukować, jak wygląda kolor, jeśli się go nie widziało!
- To ciekawe potoczne twierdzenie, muszę przyznać. Oto inne: jeśli jednocześnie beknesz, kichniesz i puścisz bąka, to umrzesz. Brzmi chyba dość wiarygodnie. Ale czy którekolwiek z nich ma jakieś naukowe podstawy?

Daniel Dennett, Słodkie sny, Prószyński i S-ka 2007

Dennett vs. *qualia*

- ▶ dlaczego zdaniem Dennetta argument Jacksona jest błędny?
 - ▶ generalny problem z *qualiami* – brak definicji
 - ▶ *if you got to ask, you ain't never gonna get to know*
 - ▶ cztery cechy charakteryzujące *qualia*
1. niewyrażalność
 2. wewnętrzność
 3. prywatność
 4. bezpośrednią dostępność

Dennett vs. *qualia*

- ▶ dlaczego zdaniem Dennetta argument Jacksona jest błędny?
- ▶ generalny problem z *qualiami* – brak definicji
 - ▶ *if you got to ask, you ain't never gonna get to know*
 - ▶ cztery cechy charakteryzujące *qualia*
 - ▶ przykład: odwrócone *qualia* a prozopagnozja i zespół Capgrasa

wstęp do filozofii umysłu i sztucznej inteligencji

21.12.2023

WYKŁAD 11

Dennett vs. *qualia*

- ▶ dlaczego zdaniem Dennetta argument Jacksona jest błędny?
- ▶ generalny problem z *qualiami* – brak definicji
 - ▶ if you got to ask, you ain't never gonna get to know
 - ▶ cztery cechy charakteryzujące *qualia*
 - ▶ przykład: odwrócone *qualia* a prozopagnozja i zespół Capgrasa
 - ▶ *qualia* odpowiadające charakterystyce Dennetta byłyby empirycznie niewykrywalne, a nie można zakładać istnienia takich bytów w teorii empirycznej
 - ▶ *qualia* są zatem co najwyżej użyteczną fikcją, ale nic w sensie naukowym im nie odpowiada

Pewnego wieczoru, przy butelce doskonałego chambertin filozof Wilfried Sellars powiedział mi: „Dan, dzięki qualiom warto żyć!”

Daniel Dennett, Dźwignie wyobraźni
i inne narzędzia do myślenia, CCP 2015

Dennett vs. *qualia*

- ▶ dlaczego zdaniem Dennetta argument Jacksona jest błędny?
- ▶ generalny problem z *qualiami* – brak definicji
 - ▶ *if you got to ask, you ain't never gonna get to know*
 - ▶ cztery cechy charakteryzujące *qualia*
 - ▶ przykład: odwrócone *qualia* a prozopagnozja i zespół Capgrasa
 - ▶ *qualia* odpowiadające charakterystyce Dennetta byłyby empirycznie niewykrywalne, a nie można zakładać istnienia takich bytów w teorii empirycznej
 - ▶ *qualia* są zatem co najwyżej użyteczną fikcją, ale nic w sensie naukowym im nie odpowiada
 - ▶ Dennett jest zatem eliminatywistą odnośnie do pojęcia *qualia*



1991



2016



Daniel C. Dennett

SŁODKIE SNY

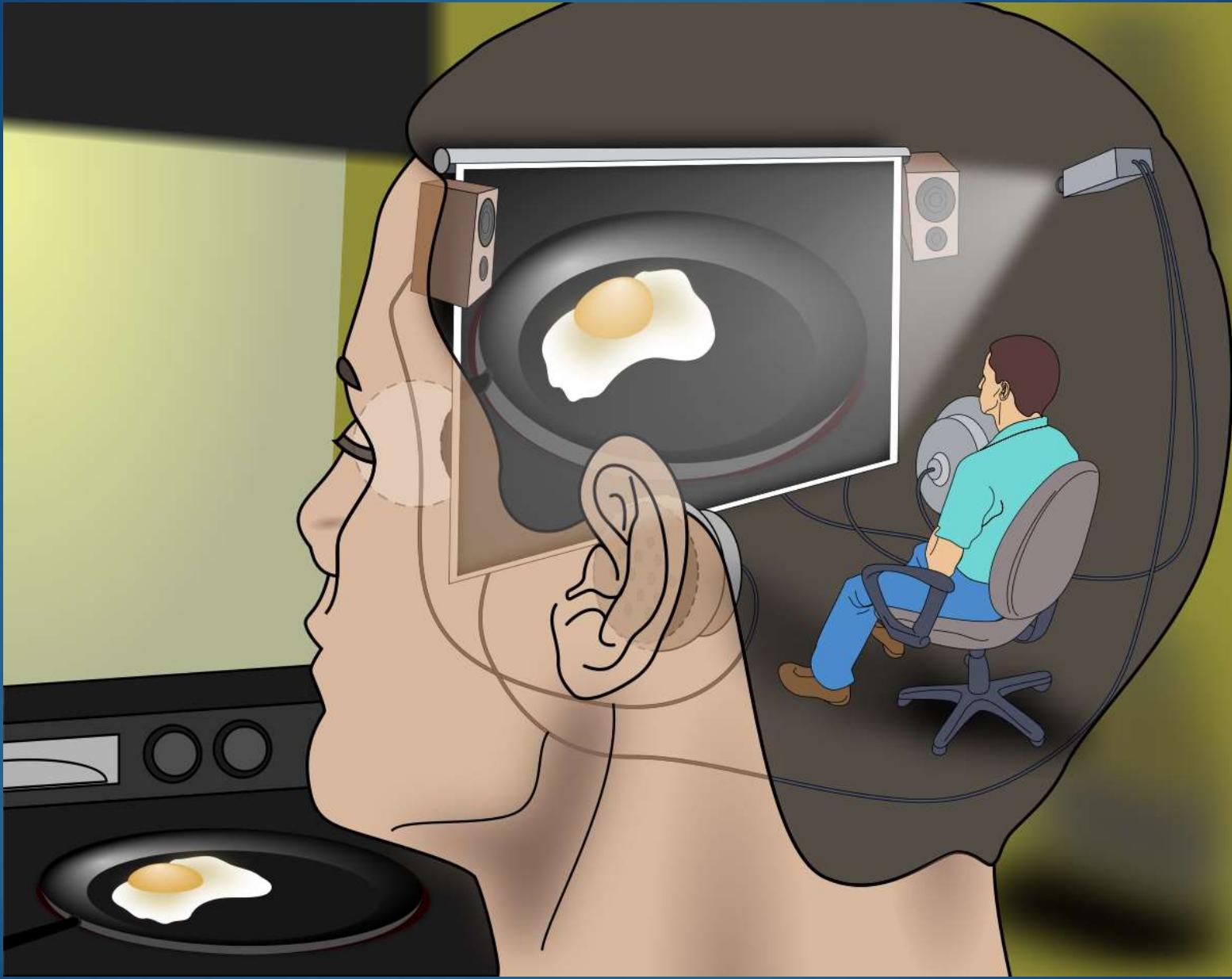
Filozoficzne
przeszkody
na drodze
do nauki
Świadomości

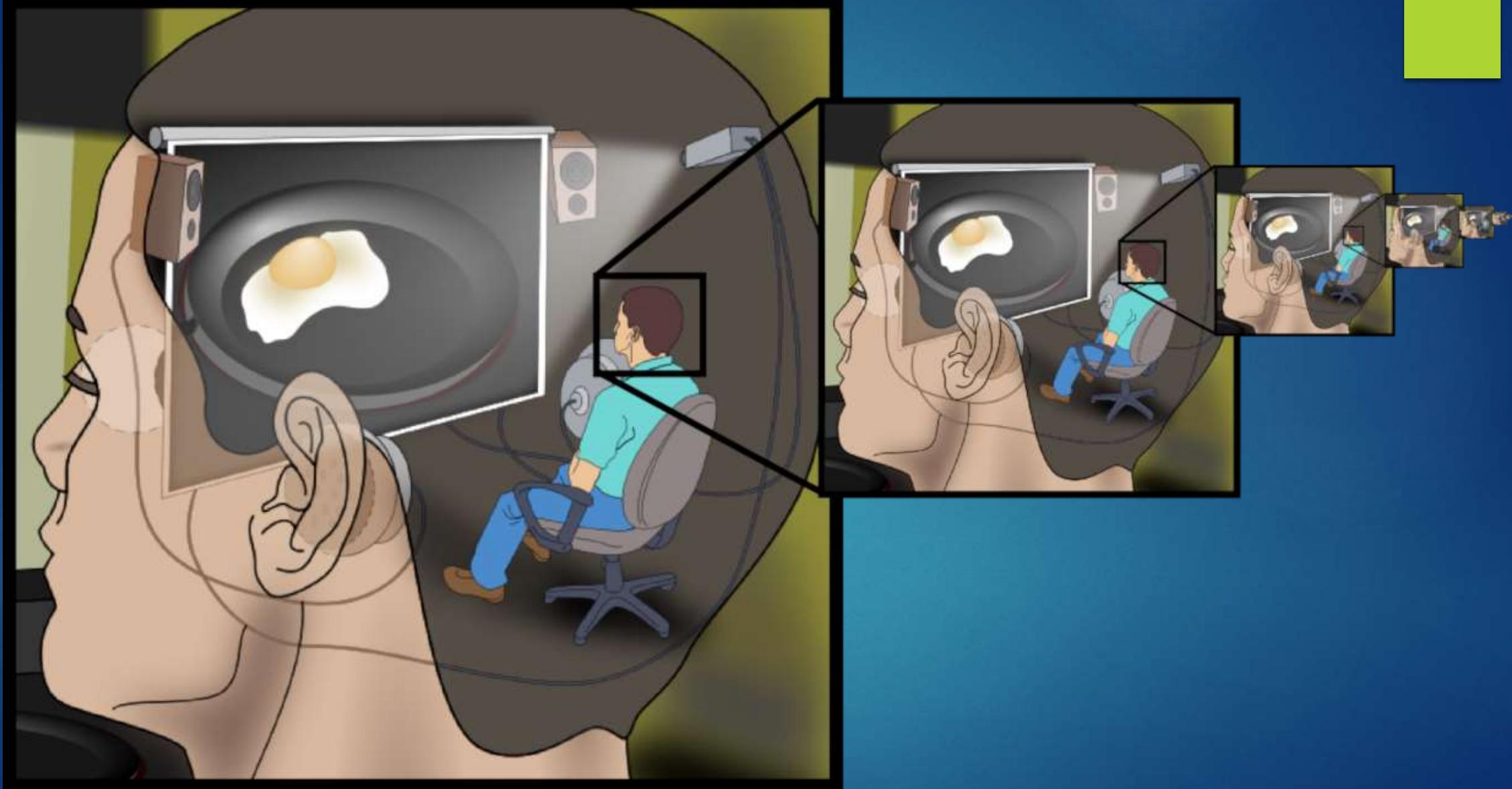


2005/2007

teatr kartezański

- ▶ większość filozoficznych koncepcji świadomości nawet jeśli zrezygnowała z kartezańskiego dualizmu, to wciąż znajduje się w pułapce kartezańskiego sposobu myślenia o świadomości
- ▶ u źródła tego sposobu myślenia leży koncepcja *punktu widzenia* w świadomym umyśle
- ▶ jest to *kartezański materializm*, współczesna pozostałość po kartezańskim dualizmie





teatr kartezański

- ▶ większość filozoficznych koncepcji świadomości nawet jeśli zrezygnowała z kartezańskiego dualizmu, to wciąż znajduje się w pułapce kartezańskiego sposobu myślenia o świadomości
- ▶ u źródła tego sposobu myślenia leży koncepcja *punktu widzenia* w świadomym umyśle
- ▶ jest to *kartezański materializm*, współczesna pozostałość po kartezańskim dualizmie
- ▶ poprawna teoria świadomości musi m.in. ścisłe polegać na empirii, nie podążać pochopnie za intuicjami, uwzględniąć w badaniu aspekt ewolucyjny

model wielokrotnych szkiców

- ▶ procesy umysłowe realizowane są w mózgu poprzez równoległe procesy interpretacji i przetwarzania informacji
- ▶ nie ma centralnego ośrodka
- ▶ informacje w mózgu poddawane są stałej „obróbce redakcyjnej” (stąd tytułowa metafora szkiców)

model wielokrotnych szkiców

Te rozproszone identyfikacje treści prowadzą z czasem do czegoś podobnego do strumienia czy ciągu narracyjnego, który można pojmować jako podlegający ciągłej redakcji przez wiele procesów rozproszonych w różnych częściach mózgu, które ciągle trwają. Ten strumień treści jedynie przypomina narrację ze względu na swoją różnorodność; w każdym momencie istnieje wiele „szkiców” fragmentów narracji na różnym etapie redakcji, w różnych miejscach mózgu. [...] Najważniejsze jest to, że model wielokrotnych szkiców unika kuszącego, choć błędnego założenia, iż musi istnieć pojedyncza narracja [...], która jest obowiązująca – czyli jest faktycznym strumieniem świadomości [...].

Daniel Dennett, Świadomość, CCP 2016



wstęp do filozofii umysłu i sztucznej inteligencji

4.01.2024

WYKŁAD 12



Colin McGinn
(ur. 1950)

misterianizm

- ▶ problemy umysł-ciało i świadomości jako *tajemnica*, zagadka nie do rozwiązania
- ▶ misterianizm umiarkowany i skrajny
- ▶ zrozumienie umysłu przekracza nasze (naturalne) zdolności poznawcze
- ▶ powód jest ewolucyjny: nasz umysł jest wynikiem doboru naturalnego
- ▶ nawet jeśli nasze przyszłe artefakty (sztuczna inteligencja) będą w stanie rozwiązać zagadkę, nie będziemy w stanie zrozumieć odpowiedzi
- ▶ introspekcja nie mówi nic o działaniu mózgu, obserwacja mózgu nie mówi nic o przeżyciach pierwszoosobowych

intencjonalność

- ▶ odnoszenie się do czegoś
- ▶ relacja stanu umysłu do jego przedmiotu



intencjonalność

- ▶ odnoszenie się do czegoś
- ▶ relacja stanu umysłu do jego przedmiotu
- ▶ *intendere arcum in* – mierzyć w coś z łuku

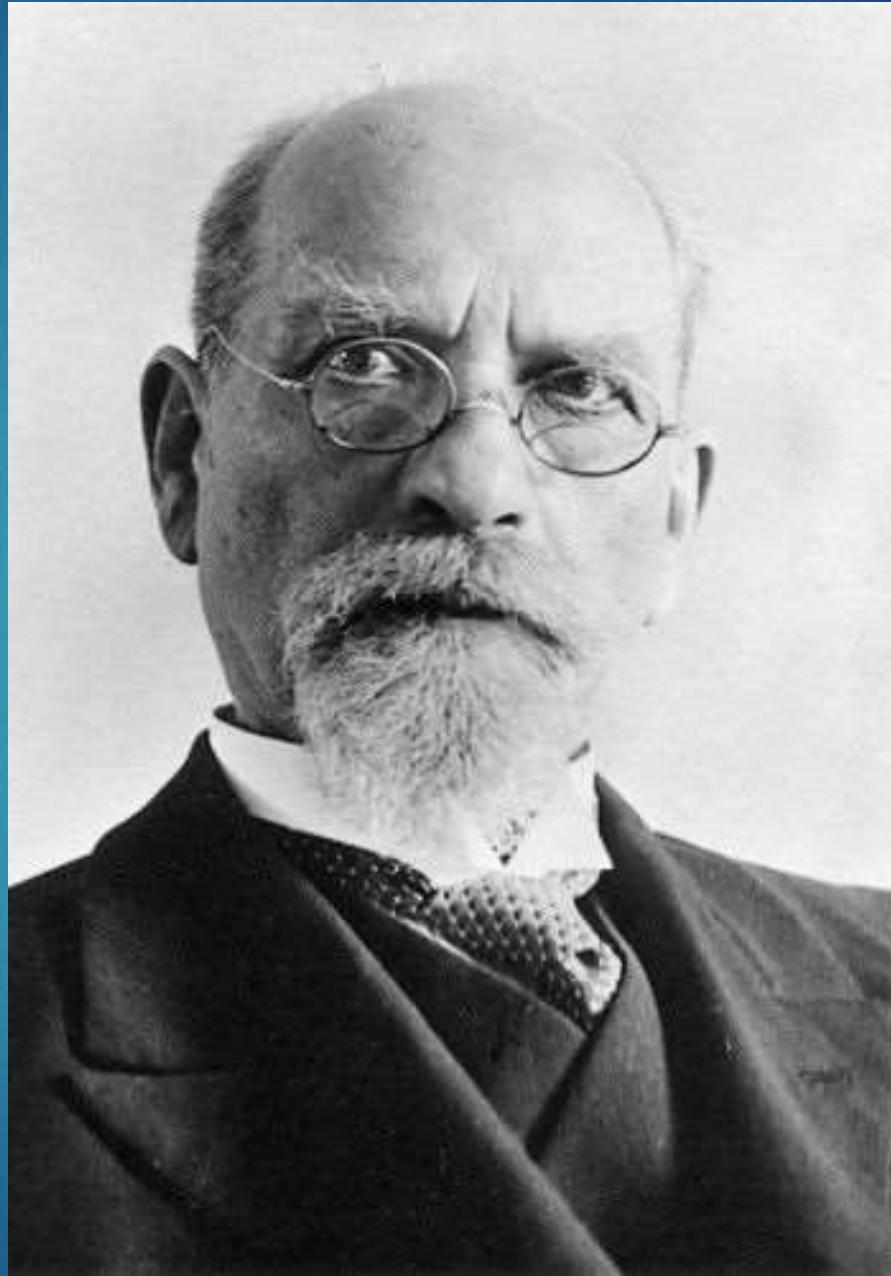
Franz Brentano
(1838-1917)



intencjonalność

- ▶ odnoszenie się do czegoś
- ▶ relacja stanu umysłu do jego przedmiotu
- ▶ *intendere arcum in* – mierzyć w coś z łuku
- ▶ intencjonalność pierwotna i wtórna
 - ▶ pierwotna dotyczy samego aktu intelektu, bezpośrednio uchwytyjącego przedmiot
 - ▶ wtórna dotyczy aktu refleksji, uchwytyjącego bezpośrednie odnoszenie się
 - ▶ intencja pierwotna dotyczy przedmiotu realnego, wtórna przedmiotu mentalnego
 - ▶ Brentano podobnie wyróżnił pierwotny i wtórny przedmiot intencjonalny
 - ▶ pierwotnym jest przedmiot odniesienia, wtórnym – akt (czynność podmiotu)
 - ▶ każdy akt psychiczny jest intencjonalny, tj. nakierowany na jakiś przedmiot

Edmund Husserl
(1859-1938)



intencjonalność

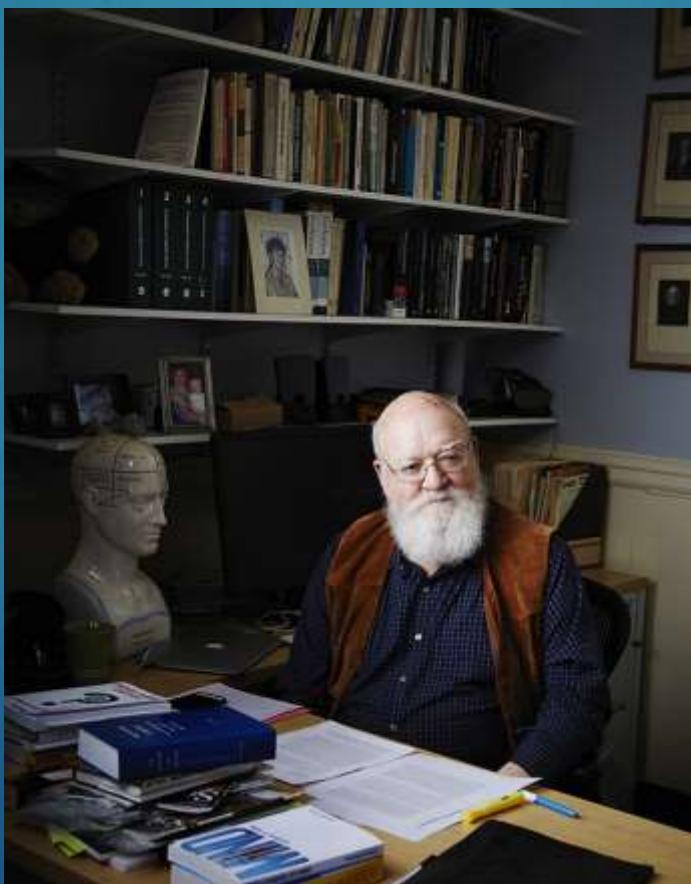
- ▶ odnoszenie się do czegoś
- ▶ relacja stanu umysłu do jego przedmiotu
- ▶ *intendere arcum in* – mierzyć w coś z łuku
- ▶ intencjonalność pierwotna i wtórna
- ▶ cechy intencjonalności według Husserla
 - ▶ dynamiczność – kształtuje jedność świadomości
 - ▶ każdy akt świadomości jest intencjonalny
 - ▶ ...ale nie każdy stan mentalny (np. czysta percepja)
 - ▶ rozróżnienie treści i przedmiotu aktów intencjonalnych
 - ▶ noemat-noeza (przedmiot i akt) – zawsze skorelowane

naturalizacja intencjonalności

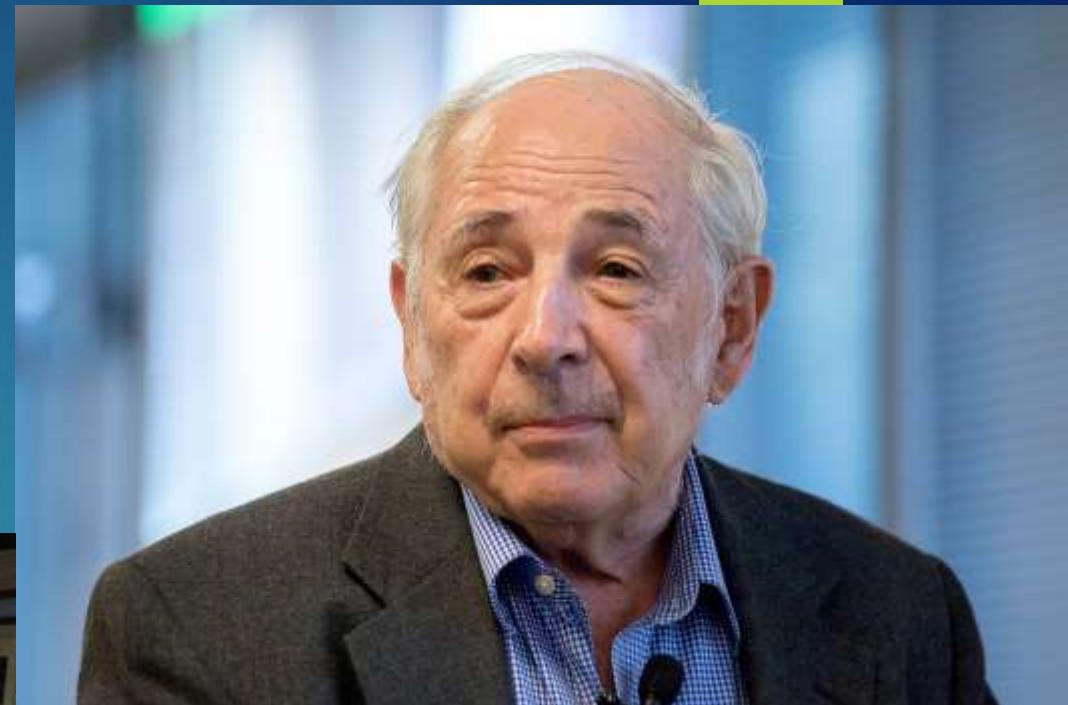




Fred Dretske
(1932-2013)



John Searle
(ur. 1932)



Daniel Dennett
(ur. 1942)

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
 - ▶ zdolność do błędnego reprezentowania
 - ▶ bycie o czymś [aboutness]
 - ▶ kształt aspektowy
 - ▶ skierowanie [directedness]

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
 - ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
 - ▶ nie wszystkie przeżycia psychiczne są intencjonalne (np. nieokreślony niepokój nie jest)
 - ▶ nie wszystkie stany intencjonalne są świadome (np. dyspozycje do przekonań)
 - ▶ niemniej intencjonalność i świadomość/umysłowość są ściśle związane
 - ▶ dwa i pół rodzaju intencjonalności: wewnętrzna, zewnętrzna, metaforyczna („jak gdyby”)
-
1. Jestem bardzo głodny.
 2. Po angielsku „I'm starving” znaczy, że jestem bardzo głodny.
 3. Muszę podpiąć kabel, bo mój komputer krzyczy, że jest głodny.

wstęp do filozofii umysłu i sztucznej inteligencji

11.01.2024

WYKŁAD 13

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
 - ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
 - ▶ nie wszystkie przeżycia psychiczne są intencjonalne (np. nieokreślony niepokój nie jest)
 - ▶ nie wszystkie stany intencjonalne są świadome (np. dyspozycje do przekonań)
 - ▶ niemniej intencjonalność i świadomość/umysłowość są ściśle związane
 - ▶ dwa i pół rodzaju intencjonalności: wewnętrzna, zewnętrzna, metaforyczna („jak gdyby”)
-
1. Jestem bardzo głodny.
 2. Po angielsku „I'm starving” znaczy, że jestem bardzo głodny.
 3. Muszę podpiąć kabel, bo mój komputer krzyczy, że jest głodny.

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
- ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
 - ▶ nie wszystkie przeżycia psychiczne są intencjonalne (np. nieokreślony niepokój nie jest)
 - ▶ nie wszystkie stany intencjonalne są świadome (np. dyspozycje do przekonań)
 - ▶ niemniej intencjonalność i świadomość/umysłowość są ściśle związane
 - ▶ dwa i pół rodzaju intencjonalności: wewnętrzna, zewnętrzna, metaforyczna („jak gdyby”)
 - ▶ intencjonalność wewnętrzna kluczowa do wyjaśnienia pozostałych
 - ▶ wyjaśnienie da się sformułować w kategoriach czysto biologicznych (przykład głodu/pragnienia)

naturalizacja intencjonalności

Oto jak działa pragnienie. Brak wody w organizmie powoduje, że nerki zaczynają wydzielać reninę, która sprawia, że z obecnego w krwiobiegu peptydu, nazywanego angiotensyną, powstaje angiotensyna 2. Substancja ta dostaje się do mózgu, do podwzgórza, czego efektem jest nasilenie się sygnałów neuronalnych w tym obszarze. To z kolei powoduje, że zwierzę świadome odczuwa pragnienie. [...] Wiemy, i jest to prosty fakt biologiczny, że pewne procesy biologiczne po prostu powodują świadome stany intencjonalne, jak pragnienie czy głód.

John Searle, Umysł, język, społeczeństwo, W.A.B. 1999

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
- ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
 - ▶ nie wszystkie przeżycia psychiczne są intencjonalne (np. nieokreślony niepokój nie jest)
 - ▶ nie wszystkie stany intencjonalne są świadome (np. dyspozycje do przekonań)
 - ▶ niemniej intencjonalność i świadomość/umysłowość są ściśle związane
 - ▶ dwa i pół rodzaju intencjonalności: wewnętrzna, zewnętrzna, metaforyczna („jak gdyby”)
 - ▶ intencjonalność wewnętrzna kluczowa do wyjaśnienia pozostałych
 - ▶ wyjaśnienie da się sformułować w kategoriach czysto biologicznych (przykład głodu/pragnienia)
 - ▶ typ i zawartość stanu intencjonalnego
 - ▶ kierunek dopasowania (umysł-świat / świat-umysł)

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
- ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
- ▶ Dennett – nastawienie intencjonalne
 - ▶ elementarna forma podmiotowości: coś nie tylko *ma skutki*, ale także *przeprowadza działania*
 - ▶ zaawansowana forma podmiotowości cechuje się ponadto intencjonalnością – a co za tym idzie świadomością

naturalizacja intencjonalności

Każda komórka [...] jest mniej więcej tak samo bezmyślna jak wirus. Czy jest możliwe, że jeśli wystarczająco wiele [z nich] połączy się ze sobą, to w rezultacie powstanie prawdziwa, świadoma osoba z autentycznym umysłem? Według współczesnej nauki nie ma innego sposobu na powstanie prawdziwej osoby. [...] jeśli nie istnieje w nas jakiś tajemniczy ekstraskładnik [...], to składamy się z robotów – lub, co oznacza dokładnie to samo, jesteśmy zbiorem trylionów makromolekularnych mechanizmów. [...] Zatem, coś, co składa się z robotów, może przejawiać autentyczną świadomość, bo jeśli cokolwiek ją przejawia, to właśnie ty.

Daniel Dennett, *Natura umysłów*, Wydawnictwo CIS 1997

naturalizacja intencjonalności

- ▶ Dretske – intencjonalność wyjaśniana w ramach reprezentacjonizmu
- ▶ Searle – intencjonalność wyjaśniana w ramach naturalizmu biologicznego
- ▶ Dennett – nastawienie intencjonalne
 - ▶ elementarna forma podmiotowości: coś nie tylko ma skutki, ale także przeprowadza działania
 - ▶ zaawansowana forma podmiotowości cechuje się ponadto intencjonalnością – a co za tym idzie świadomością
 - ▶ nastawienie fizyczne, nastawienie projektowe, nastawienie intencjonalne
 - ▶ psychologia potoczna pociąga za sobą przyjęcie nastawienia intencjonalnego
 - ▶ system intencjonalny to cokolwiek, wobec czego przyjmujemy nastawienie intencjonalne
 - ▶ analiza systemów intencjonalnych pozwala ostatecznie zasypać przepaść między psychologią potoczną a działaniem organizmu jako systemu fizycznego

wolna wola

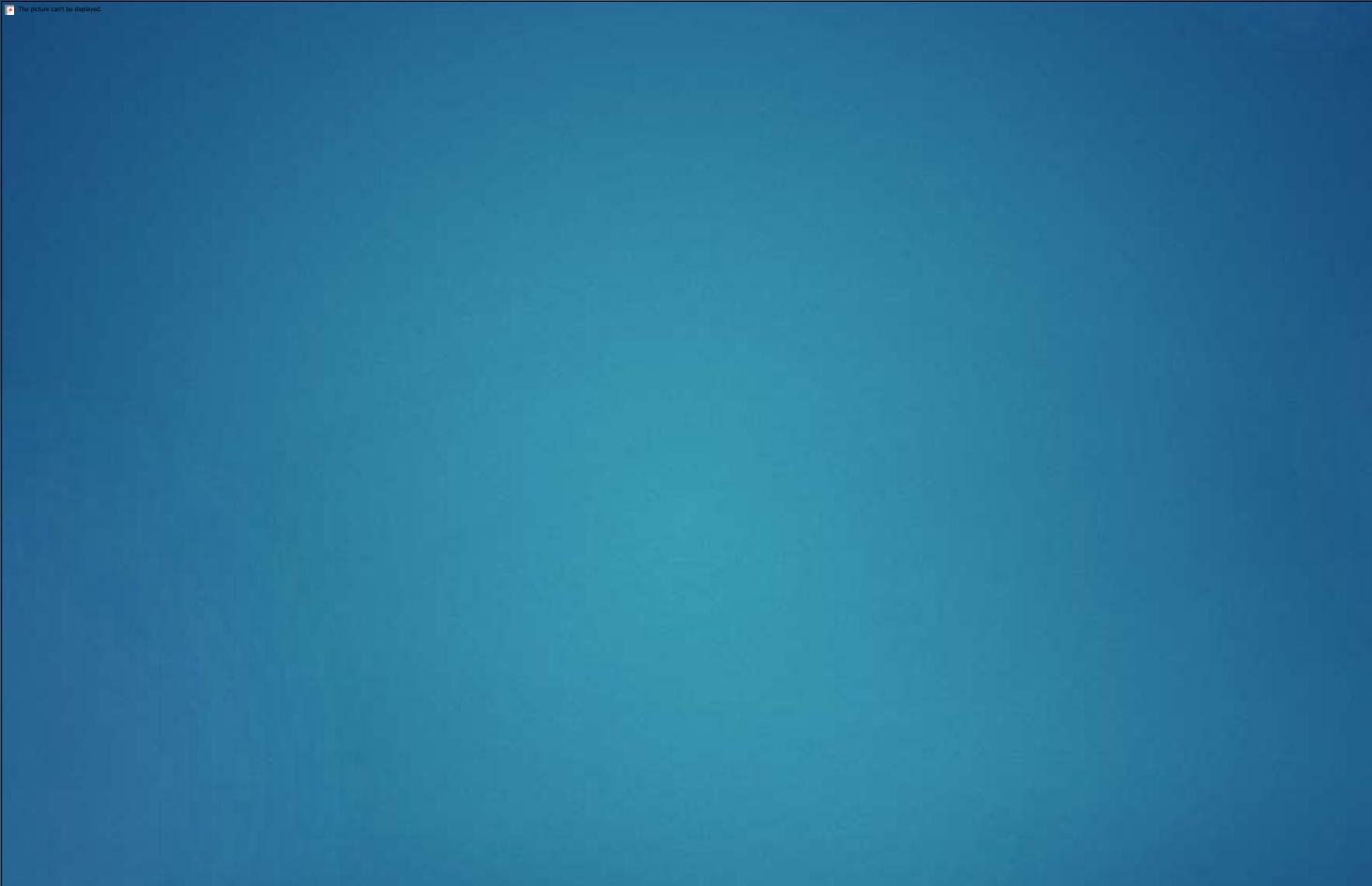
- ▶ co to tak naprawdę znaczy mieć wolną wolę?
 - ▶ robienie tego, co chce się robić
 - ▶ tego co chce się robić w ramach dostępnych możliwości
 - ▶ dyspozycja (zdolność) do czynienia danej rzeczy: jeśli w danej sytuacji pewna osoba zdecydowałaby się postąpić inaczej, miałaby możliwość postąpić inaczej

wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
 - ▶ robienie tego, co chce się robić
 - ▶ tego co chce się robić w ramach dostępnych możliwości
 - ▶ dyspozycja (zdolność) do czynienia danej rzeczy
 - ▶ środowisko umożliwiające realizację dyspozycji

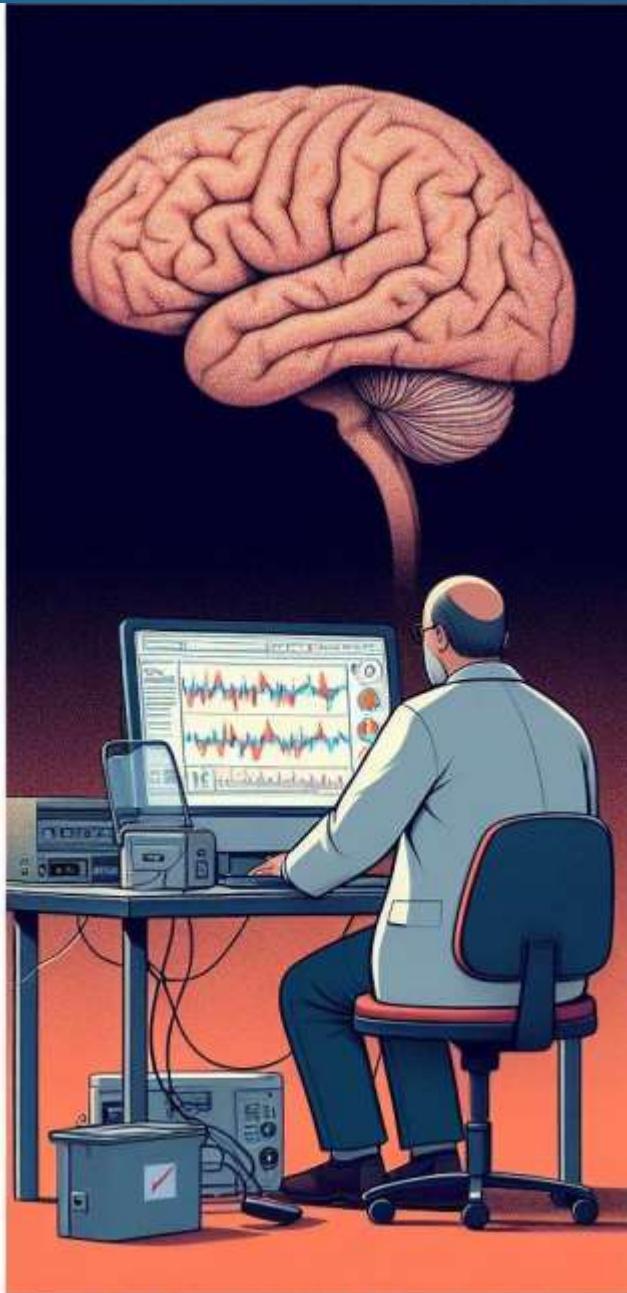
wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
 - ▶ przy założeniu (jakiejś formy) naturalizmu, podmiot wolnej woli ma charakter czysto fizyczny, podlega zatem prawom fizyki
 - ▶ jeśli podmiot podlega prawom fizyki, to działa w pełni deterministycznie (zasada domknięcia przyczynowego)
 - ▶ jednocześnie bez przyjęcia wolnej woli podmiotu niełatwo wyjaśnić niektóre jego zachowania
 - ▶ ponadto przyjęcie wolnej woli wydaje się niezbędne do przypisania komuś odpowiedzialności moralnej



Harry Frankfurt
(1929-2023)





wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
 - ▶ przy założeniu (jakiejś formy) naturalizmu, podmiot wolnej woli ma charakter czysto fizyczny, podlega zatem prawom fizyki
 - ▶ jeśli podmiot podlega prawom fizyki, to działa w pełni deterministycznie (zasada domknięcia przyczynowego)
 - ▶ jednocześnie bez przyjęcia wolnej woli podmiotu niełatwo wyjaśnić niektóre jego zachowania
 - ▶ ponadto przyjęcie wolnej woli wydaje się niezbędne do przypisania komuś odpowiedzialności moralnej
 - ▶ o wolności i odpowiedzialności nie decyduje, czy podmiot miał możliwość uczynić inaczej, ale czy jest źródłem swych działań

wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
- ▶ kompatybilizm i inkompatybilizm
 - ▶ oba stanowiska przyjmują determinizm
 - ▶ zdaniem inkompatybilistów wolna wola jest nie do pogodzenia z determinizmem

P1. Jeśli ktoś ma wolną wolę, to może postąpić inaczej niż postępuje

P2. Jeśli determinizm jest prawdziwy, to nikt nie może postąpić inaczej niż postępuje

W. Zatem jeśli determinizm jest prawdziwy, to nikt nie ma wolnej woli

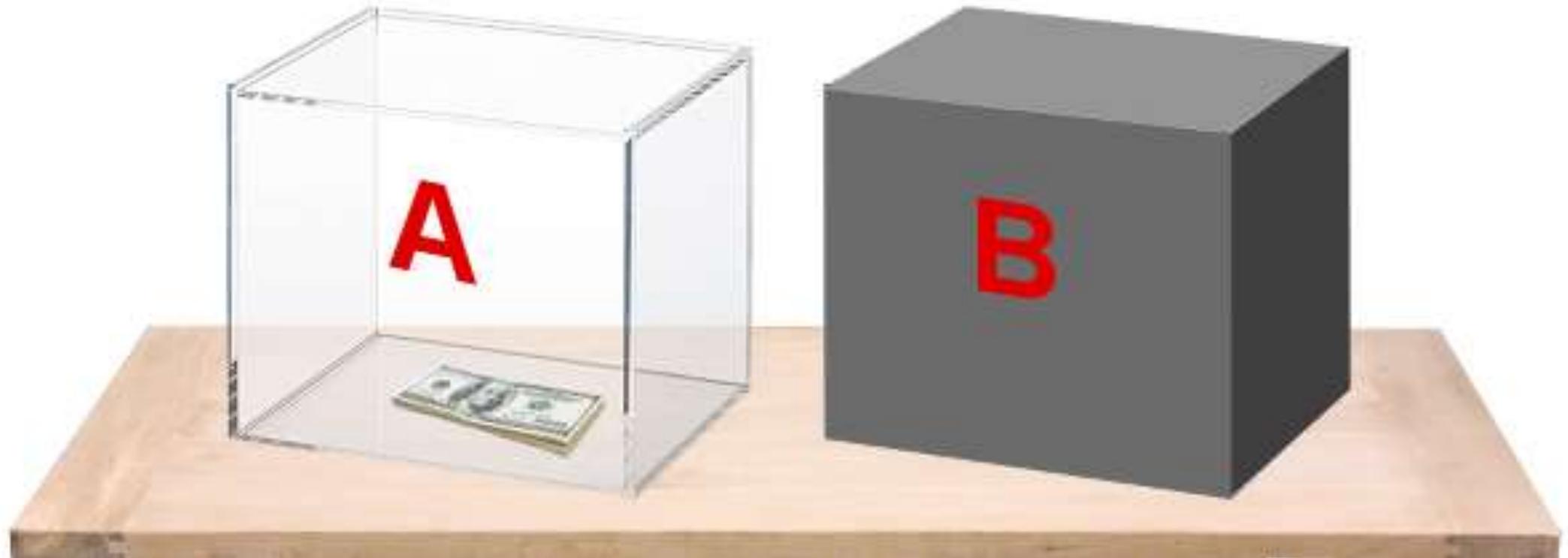
wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
- ▶ kompatybilizm i inkompatybilizm
 - ▶ oba stanowiska przyjmują determinizm
 - ▶ zdaniem inkompatybilistów wolna wola jest nie do pogodzenia z determinizmem

P1. Jeśli ktoś ma wolną wolę, to jest źródłem swoich działań

P2. Jeśli determinizm jest prawdziwy, to nikt nie jest ostatecznym źródłem swoich działań

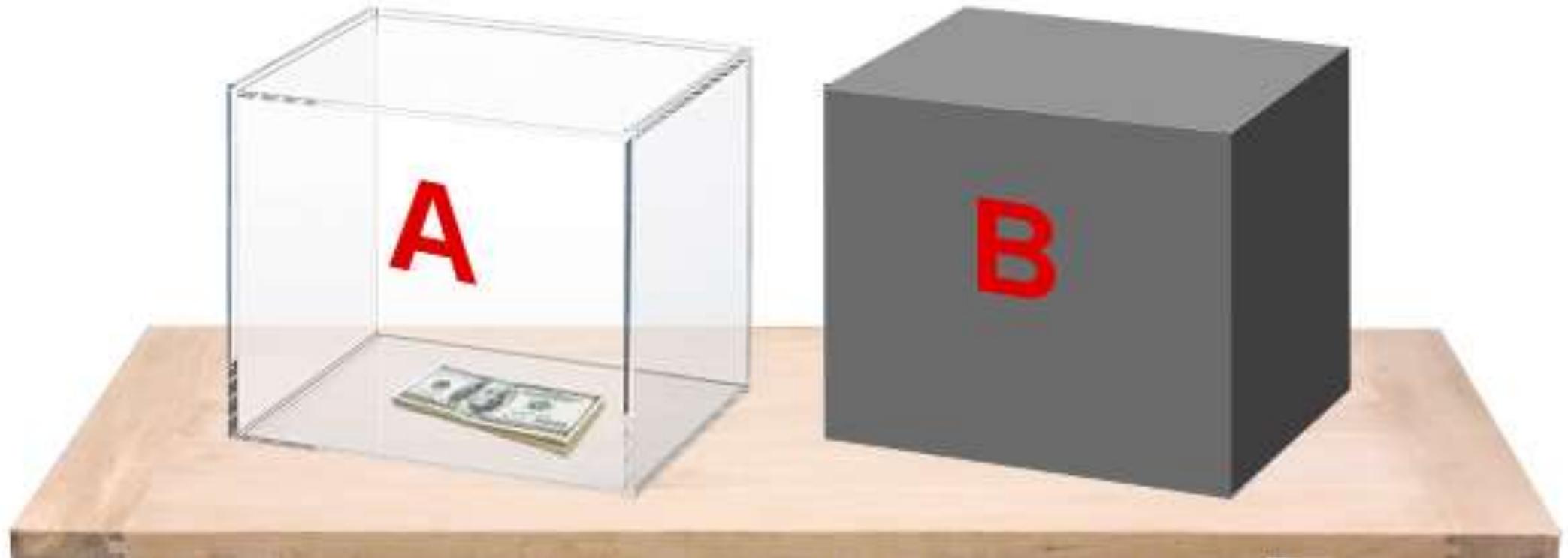
W. Zatem jeśli determinizm jest prawdziwy, to nikt nie ma wolnej woli



wstęp do filozofii umysłu i sztucznej inteligencji

18.01.2024

WYKŁAD 13



The original post was viewed more than 200,000 times, with well over a thousand comments. We tallied 31,854 votes before we closed submissions. And the results are:

- I choose box B: **53.5 per cent**
- I choose both boxes: **46.5 per cent**

wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
- ▶ kompatybilizm i inkompatybilizm
 - ▶ oba stanowiska przyjmują determinizm
 - ▶ zdaniem inkompatybilistów wolna wola jest nie do pogodzenia z determinizmem
 - ▶ zdaniem kompatybilistów wolna wola jest do pogodzenia z determinizmem
 - ▶ Frankfurt: pragnienia pierwszego i drugiego rzędu
 - ▶ Chalmers: wolność woli i wolność wolności woli



I always know what I'm going to do
before I do it. OK?



I did it because I knew
I was going to do it.

wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
- ▶ kompatybilizm i inkompatybilizm
 - ▶ oba stanowiska przyjmują determinizm
 - ▶ zdaniem inkompatybilistów wolna wola jest nie do pogodzenia z determinizmem
 - ▶ zdaniem kompatybilistów wolna wola jest do pogodzenia z determinizmem
 - ▶ Frankfurt: pragnienia pierwszego i drugiego rzędu
 - ▶ Chalmers: wolność woli i wolność wolności woli
 - ▶ Dennett: dwie loterie

wolna wola

argument inkompatybilistyczny Galena Strawsona

P₁. W każdej sytuacji robisz to, co robisz, ponieważ jesteś, jaki jesteś

P₂. Aby być ostatecznie odpowiedzialnym za to, co robisz, musisz być ostatecznie odpowiedzialny za to, jaki jesteś

P₃. Nie możesz być ostatecznie odpowiedzialny za to, jaki jesteś

W. Nie możesz być ostatecznie odpowiedzialny za to, co robisz

wolna wola

argument antyssaczy Dennetta

P₁. Ssak jest ssakiem, ponieważ jest taki, jaki jest

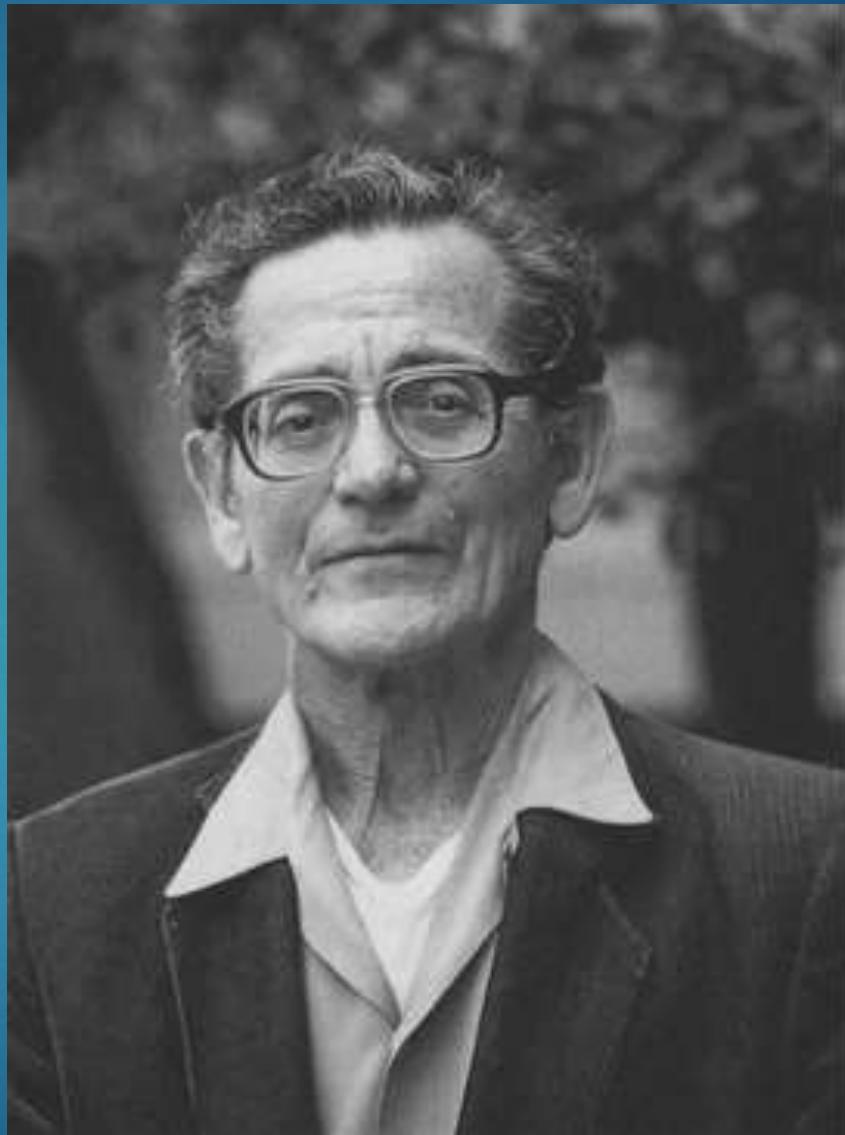
P₂. Aby być ssakiem, musisz się stać taki, jaki jesteś, mając matkę będącą ssakiem

P₃. To samo musi być prawdziwe w odniesieniu do twojej matki, jej matki i ad infinitum, co jest niemożliwe

W. Nie możesz być ssakiem, ponieważ ssaki są niemożliwe



Benjamin Libet
(1916-2007)



wolna wola

- ▶ co to tak naprawdę znaczy mieć wolną wolę?
- ▶ problem wolnej woli w filozofii umysłu
- ▶ kompatybilizm i inkompatybilizm
- ▶ badania Libeta
 - ▶ dotyczyło następstwa czasowego między świadomym podjęciem działania (ruchem ręki) a początkiem aktywności obszarów mózgu odpowiedzialnych za ten ruch
 - ▶ symetryczny potencjał gotowości (SPG)
 - ▶ SPG pojawiał się średnio 350 milisekund przed podjęciem świadomej decyzji
 - ▶ krytyka metodologii i interpretacji wyników
 - ▶ problem z właściwym modelem SPG

tożsamość osobowa

- ▶ ogólny problem: co decyduje o ciągłości naszej jaźni?

t_1

.....

t_2



 t_1  t_2

tożsamość osobowa

- ▶ ogólny problem: co decyduje o ciągłości naszej jaźni?
- ▶ tym samym jest to pytanie o to, czym jest jaźń (czym jest osoba)
 - ▶ właściwości istotowe (esencjalne) i przypadłościowe (akcydentalne)

tożsamość osobowa

- ▶ ogólny problem: co decyduje o ciągłości naszej jaźni?
- ▶ tym samym jest to pytanie o to, czym jest jaźń (czym jest osoba)
- ▶ różne wymiary tożsamości
 - ▶ tożsamość diachroniczna (w chwili t_1)
 - ▶ tożsamość synchroniczna (w przedziale t_1-t_2)
 - ▶ tożsamość osobowości (ta sama, ale nie taka sama) → stopniowalna
 - ▶ tożsamość osoby (tożsamość osoby w czasie t_1 z osobą w czasie t_2) → zero-jedynkowa

tożsamość osobowa

- ▶ ogólny problem: co decyduje o ciągłości naszej jaźni?
- ▶ tym samym jest to pytanie o to, czym jest jaźń (czym jest osoba)
- ▶ różne wymiary tożsamości
- ▶ dwa zasadnicze punkty widzenia
 - ▶ tożsamość jako właściwość prosta
 - ▶ źródło w filozoficznym racjonalizmie, zwłaszcza dualizmie Kartezjusza
 - ▶ istnieje niezmienna jaźń, „twarde jądro” naszej osobowości
 - ▶ jaźń nie ma charakteru empirycznego, ale metafizyczny

tożsamość osobowa

- ▶ ogólny problem: co decyduje o ciągłości naszej jaźni?
- ▶ tym samym jest to pytanie o to, czym jest jaźń (czym jest osoba)
- ▶ różne wymiary tożsamości
- ▶ dwa zasadnicze punkty widzenia
 - ▶ tożsamość jako właściwość złożona
 - ▶ źródło w filozoficznym empiryzmie, zwłaszcza koncepcjach Locke'a i Hume'a
 - ▶ kluczowa jest ciąłość tożsamości, redukowana do faktów fizycznych i psychicznych
 - ▶ jaźń ma charakter *stricto empiryczny*

tożsamość osobowa

- ▶ teorie ciągłości psychicznej

t_1

.....

t_2



$t_1 \rightarrow t_x \rightarrow t_{x+1} \rightarrow t_{x+2} \rightarrow t_{x+3} \rightarrow t_{x+4} \rightarrow t_2$

tożsamość osobowa

- ▶ teorie ciągłości psychicznej
 - ▶ ciągłość wspomnień

The picture can't be displayed.

John Locke
(1632-1704)

tożsamość osobowa

- ▶ teorie ciągłości psychicznej
 - ▶ ciągłość wspomnień
 - ▶ wiązkowa teoria jaźni



David Hume
(1711-1776)

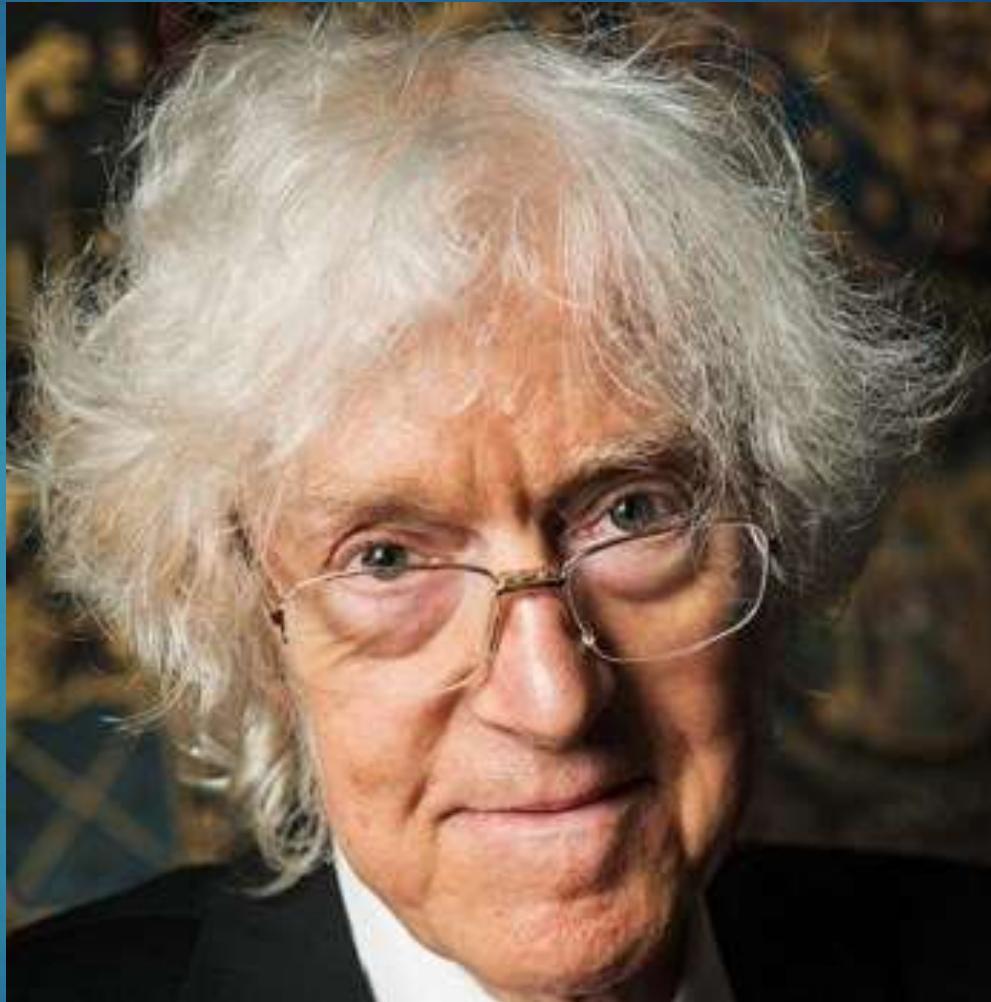


tożsamość osobowa

- ▶ teorie ciągłości psychicznej
 - ▶ ciągłość wspomnień
 - ▶ wiązkowa teoria jaźni
 - ▶ wersje współczesne



Derek Parfit
(1942-2017)



tożsamość osobowa

- ▶ teoria relacji R
 - ▶ ciągłość wyznaczona jest przez odpowiednią relację
 - ▶ zamiast o tożsamości, można mówić o przetrwaniu
 - ▶ jeśli ciągłość fizyczna i psychiczna są tylko częściowe, czasami nie da się jednoznacznie udzielić odpowiedzi na pytanie o tożsamość osobową
 - ▶ kryterium ciągłości fizycznej: czasoprzestrzenna ciągłość mózgu i ciała
 - ▶ kryterium ciągłości psychicznej: pamięć przeżyciowa
 - ▶ psychiczne powiązanie i psychiczna ciągłość
 - ▶ to pierwsze wtedy, gdy przypominamy sobie przeszłe przeżycia
 - ▶ to drugie, gdy przy braku takiego powiązania istnieje powiązanie przez zachodzący na siebie szereg wspomnień