

FILOZOFIA INFORMACJI, UMYSŁU, ŚWIADOMOŚCI I WOLI

Lukasz Lazarz

Słowa kluczowe: *Filozofia, filozofia umysłu, kognitywistyka, filozofia sztucznej inteligencji, filozofia psychologii, filozofia ekonomii, etyka, estetyka, informacja, umysł, świadomość, wiedza, emocje, wrażenia, wolna wola, szczęście, piękno, dobro, wartość.*

Abstrakt:

Na wstępie chciałbym wyraźnie zaznaczyć, że filozoficzny styl niniejszej pracy nawiązuje do słów Sthephena Hawkinga, które wypowiedział na konferencji w Google w 2011 roku, a które mogłyby stać się jej mottem: „Filozofia umarła”. Oczywiście nie chodzi zapewne autorowi tych słów o to, że umarła możliwość zadawania pytań, także najbardziej ogólnych. Bardziej chodziło mu zapewne o tym, że nauka dotarła już w swoich dziedzinach do pewnych granic poznania, które każą zadawać pytania o to co jest poza nimi, przy pomocy bardziej precyzyjnego języka do opisu świata. „Klasyczna” filozofia wobec postępu nauk ścisłych i przyrodniczych staje się jedynie pomocą, czasami pewną ramą pojęciową dla problemów rozważanych na gruncie meta nauki. Można powiedzieć, że ontologię zastępuje filozofii przyrody, epistemologię filozofia umysłu szczególnie ta czerpiąca wprost z kognitywistyki, eudajmonię i aksjologię zastępuje meta refleksja psychologiczna czy ekonomiczna. Można też jednak stwierdzić, że te meta refleksja nad nauką i filozofia od swoich początków wzajemnie się przeplatają a dobra filozofia po prostu musi być uprawiana w bardzo ścisłych związkach z nauką. **W niniejszej pracy, poza wstępnymi uwagami ontologicznymi związanymi z fundamentalnym dla tej pracy pojęciem informacji, chciałbym przede wszystkim omówić problem Umysłu, Świadomości, Woli (i innych przyczyn ludzkiego postępowania). Czyli innymi, prostymi słowy; poza wstępnymi uwagami na temat tego tego (I) jaki jest ten Świat (z czego jest zbudowany), chciałbym wskazać (II) kim z punktu współczesnej nauki o Umyśle jesteście, czyli jak postrzegamy i czujemy Świat oraz (III) jak postępujemy i jak powinniśmy postępować (lub jak powinno być).**

0. Wstęp:

Powstanie nauk umożliwiających podjęcie bardziej naukowej analizie tych fundamentalnych problemów ściśle związane było z pierwszą rewolucją przemysłową, która rozpoczęła się pod koniec XVIII wieku. Kryzys energetyczny spowodowany brakiem drewna zmusił mieszkańców Anglii do poszukiwania innego źródła energii cieplnej. Okazały się nim znaczne zasoby łatwo dostępnego węgla, położone w pobliżu rzek, którego transport okazał się tani i łatwy. Tak zwana druga rewolucja przemysłowa przypada na drugą połowę XIX i początek XX stulecia. Spowodowana została gwałtownym rozwojem nauki, któremu towarzyszyło powstanie nowych rozwiązań technicznych (np. Żarówka w 1879 r.). W zakresie nauk biologicznych przełomowe znaczenie miały powstałe teorie Charlesa R. Darwina. W ogłoszonych dziełach naukowych: „O powstaniu gatunków” (1859 r.) i „O pochodzeniu człowieka” (1871 r.) dowodził, że świat roślinny i zwierzęcy powstał drogą stopniowego rozwoju i przemian organizmów prostych w coraz bardziej złożone i różnorodne. Albert Einstein (1905 r.) ogłosił teorię względności, bezprecedensowo modyfikującej pojmowanie takich pojęć, jak przestrzeń, czy czas. Narodziły się: fizyka i mechanika kwantowa. W zakresie psychologii i psychiatrii coraz częściej zajmowano się człowiekiem za pomocą metody doświadczalnej w oderwaniu od jego pojęcia duszy. Powstała psychologia fizjologiczna i behawioryzm. W zakresie metamatematyki prace Russela, Hilberta, Godla, Churcha i Turinga pokazały ograniczanie formalizacji opisu. Alan Turing (1936 r.) przebywając w Cambridge Turing napisał swoją prawdopodobnie najważniejszą pracę matematyczną „On Computable Numbers”, w której wprowadził abstrakcyjną maszynę, która była w stanie wykonywać zaprogramowaną matematyczną operację, czyli algorytm. Równoległe do problemów logicznych, matematycy zajmowali się także przetwarzaniem informacji bardziej pod kątem zastosowań praktycznych związanych czy to z transmisją i kompresją danych. Claude Shannon i Warren Weaver opracowali matematyczną teorię informacji.

Równoległe rozwijała się inżynieria a z nią technologia komputerowa. Prace Babbage’a, Ady Lovelace, Konrada Zeusa doprowadziły do powstania pierwszych prototypów komputerów. W USA powstały komputer Atanasoffa-Berry’ego (krótko ABC), ENIAC Eckerta i Mauchly’ego. Zamiast zwykłych przekaźników dysponowały lampami elektronowymi.

Wszystkie te zmiany umożliwiły powstanie cybernetyki, która zajmowała się opisywaniem układów biologicznych jako układów obliczeniowych. To utorowało z kolei powstaniu takich dyscyplin jako kognitywistyka, psychologia poznawcza czy sztuczna inteligencja. W 1956 r. roku w MIT odbyły się warsztaty poświęcone zagadnieniu informacji oraz odbyły się warsztaty w Dartmouth, które w zamiarze miały doprowadzić do rozwoju sformalizowanych form myślenia (Claude Shannon, Marvin Minsky i John McCarthy). Następnie rozwój neurobiologii oraz prace nad modelami neuronów (Warren McCulloch, Walter Pitts) oraz układami równoległe przetwarzającymi informacje (Donald Hebb, Frank Rosenblatt, John Hopfield) ostatecznie ukształtowały współczesny obraz nauki o umyśle. W ostatnich latach niesłychany postęp technologii sztucznej inteligencji spowodował, że za jej pomocą wyniki tych nauk szczególnie mocno zaczęły docierać do najszerzej publiczności (Deep Mind, AlfaGo, ChatGP

SPIS TREŚCI

<p style="text-align: center;">1. INFORMACJA</p> <p><i>czyli z czego zbudowany JEST ŚWIAT?</i></p> <p>Filozofia nauki, ontologia, metafizyka</p>	<p style="text-align: center;">2. UMYŚŁ, ŚWIADOMOŚĆ</p> <p><i>czyli kim JESTEM, jak POSTRZEGAM i CZUJĘ?</i></p> <p>Filozofia umysłu, epistemologia, metafizyka</p>	<p style="text-align: center;">3. SZCZĘŚCIE, (PIĘKNO), DOBRO i WARTOŚĆ</p> <p><i>czyli jak POSTĘPUJĘ lub jak POWINIENEM postępować lub jak POWINNO BYĆ?</i></p> <p>Aksjologia, Eudajmonia, Etyka, Estetyka, Metafizyka</p>
<p>/</p> <p>1.1. Materia (Fizyka)</p>	<p>2.1. Umysł, wiedza, myśl, myślenie (Kognitywistyka, sztuczna inteligencja)</p>	<p>2.6. Wartość, dobro, normatywność, użyteczność (Filozofia)</p>
Znane		Poznanie
<p>1.2. Energia (Fizyka)</p>	<p>2.2. Emocje (Psychologia, kognitywistyka)</p>	<p>2.5. Szczęście, dobrostan psychiczny, użyteczność (Filozofia, psychologia, ekonomia)</p>
Nieznane		Czucie
<p>1.3. Informacja (Filozofia, matematyka, kognitywistyka)</p>	<p>2.3. Wrażenia, Świadomość i Umysł (Filozofia, neurobiologia, kognitywistyka)</p>	<p>2.4. Jaźń, wolna wola, instynkt, intencja (Filozofia umysłu, neuropsychologia, psychologia)</p>

1. Informacja czyli z czego zbudowany jest Świat?

1.1. Materia

Pojęcie "materii" jest jedną z najstarszych kategorii, jakie służyły racjonalnemu wyjaśnieniu Świata. Na rozumienie materii jako tworzywa, z jakiego "zbudowane" są poszczególne rzeczy wskazuje już sam Źródłosłów łacińskiego terminu materia, który w jednym ze swoich podstawowych znaczeń odnosił się do materiału, z jakiego coś można zrobić, budulca, tworzywa, substancji - w fizycznym sensie tego słowa. Termin ten uważany jest za najbliższy greckiemu „hyle”, którego użycie w znaczeniu "tworzywa" rzeczy materialnych utrwaliło się w filozofii za sprawą Arystotelesa.

Spór co do zakresu znaczeniowego pojęcia „materia” pojawił się wśród fizyków już w epoce formowania się fizyki jako nauki, tj. w wieku XVII. Według Kartezjusza, materia to byt rozciągły w przestrzeni, wypełniający ją całą, przy braku próżni absolutnej, podzielny nieskończenie, nie tylko w postaci ciał (cząstek), ale subtelny materiał przenoszący oddziaływanie między nimi.¹ Ze względu na swój całkowicie spekulatywny jak na ten czas charakter zwyciężyła koncepcja przedstawiona przez Izaaka Newtona: świat składa się z ciał zbudowanych z twardych i sztywnych, niepodzielnych cząstek, nieprzenikliwych (jedno miejsce może zajmować jedna tylko cząstka), rozciągniętych w absolutnej przestrzeni i trwających w absolutnym czasie, zdolnych do ruchu, obdarzonych bezwładnością (masą bezwładną) i ciężkością (masą ważką), rozdzielonych absolutną próżnią; oddziaływania przenoszą się na odległość (przez próżnię, bez pośrednictwa ciał materialnych), z nieskończoną prędkością.² Podstawowym atrybutem materii i miarą jej ilości (a w fizyce posiadanie cech mierzalnych decyduje o uznaniu danego bytu za realny), jest w tym ujęciu masa. Koncepcja ta wydawała się mieć mocne oparcie w ścisłych prawach fizyki newtonowskiej (dynamiki i teorii grawitacji) i stała się filozoficzną podstawą dalszego rozwoju mechaniki klasycznej, a następnie, po uznaniu ciepła za skutek ruchu cząstek (Herapath, Carnot), kinetycznej teorii gazów i termodynamiki. Rozwój fizyki wydawał się dowodzić nieistnienia imponderabiliów – bytów istniejących obiektywnie, rozciągniętych, lecz nie posiadających masy, takich jak ciepłiki.

1.2. Energia

Pojęcie energii wyraża ogólnie aktywność. Pierwotnie nie dotyczyło ono jedynie fizyki, ale dotyczyło także ludzkiego ducha i ludzkiej woli. Tak rozumiany filozoficznie termin (od gr. „ἐνέργεια”) był od dawna używany na oznaczenie działania, aktu, entelechii czy siły. W filozofii średniowiecznej wyrażano go łacińskimi: virtus lub vis. Współcześnie „energia” jest pojęciem przynależnym fizyce. Oznacza ono zdolność jakiejś siły do wykonania pracy. O ile pojęcia masy i siły występują w Principiach Newtona w całkowicie poprawnym znaczeniu, znaczeniu tożsamym ze znaczeniem współczesnym, o tyle pojęcie „energia” tam nie występuje. Jednak twórcy mechaniki zdawali sobie sprawę z zachowania całkowitej energii w procesach mechanicznych (teoria zegara wahadłowego Huygensa jest implicite oparta na całej energii). Kartezjusz wypowiedział myśl, że w procesach przyrodniczych pewna wielkość powinna być stała w czasie. Za taką uznał iloczyn masy i prędkości, który nazwał „ilością ruchu”. Leibniz wypowiedział taką samą myśl, ale za wielkość stałą w czasie uznał iloczyn masy i kwadratu prędkości, który nazwał „siłą żywą” (vis viva). Dało to początek sławetnej debacie o tym, co jest prawdziwą miarą „ilości ruchu”. Równoległe do idei stałości w czasie pewnej wielkości kształtowała się dzięki Leibnizowi idea, że procesy fizyczne odbywają się zgodnie z

¹ Kartezjusz, Meteorologia, 1637, Zasady filozofii, 1644.

² Izaak Newton, Philosophiae naturalis principia mathematica 1687, Optyka 1704.

pewną zasadą najmniejszego oporu. Problemy ze sformułowaniem współczesnego rozumienia zasady zachowania energii mogły zostać rozwiązane wraz z rozwojem nauki o ciepłe.

Za twórców współczesnego rozumienia prawa zachowania energii uważa się dość powszechnie Mayera, Helmholtza i Joule'a. Joule wyznaczył mechaniczny równoważnik ciepła tzn. ilość ciepła, którą można uzyskać z danej ilości energii mechanicznej. Pod wpływem filozoficznych prądów poszukiwania teorii unitarnych w XIX w. pojęcie energii zaczęło odgrywać fundamentalną rolę także w rozwijającej się nauce. Udowodniano, jak można zamieniać jedną formę energii w inną, a tych form energii poznawano coraz więcej: od energii mechanicznej i cieplnej poprzez energię chemiczną, akustyczną, światła, magnetyczną, elektryczną itd. Starano się jednak sprowadzić tę rozbudowującą się klasyfikację do dwóch form podstawowych: energii potencjalnej i kinetycznej. W połowie XIX w. zasada zachowania energii zaczęła przybierać dwie formy wyrazu: „w układzie izolowanym energia całkowita jest stała” lub „energii nie można wytworzyć ani zniszczyć”. W ten sposób fizyka przybrała postać nauki nie tylko o materii, ale i o energii. Na dodatek w Szczególnej Teorii Względności Albert Einstein dodatkowo wykazał pewne związki między masą i energią w postaci słynnego równania $E=mc^2$.

1.3. Informacja

Teoria informacji ma swoje źródło w fizyce statystycznej, konkretnie w pojęciu entropii. Entropia jest miarą stopnia nieuporządkowania układu, czyli także miarą naszej o nim niewiedzy. Informacja natomiast jest zawsze związana ze zmniejszaniem naszej niewiedzy. W 1948 r. Claude Shannon opublikował artykuł „Mathematical Theory of Communication”³, w którym przedstawił matematyczny model przesyłania informacji za pomocą środków komunikacji, takich jak telegraf czy nadajnik radiowy. Wprowadził on także miarę informacji; funkcję entropii H , która spełnia wymienione warunki i definiuje ją za pomocą wzoru:

$$S = -k \sum p(i) \ln(p(i)).$$

Matematyczna teoria komunikacji nie zajmuje się ani strukturalnymi cechami informacji, ani kwestiami dotyczącymi znaczenia informacji, ani nie odnosi się ona też do żadnej wyraźnie sprecyzowanej definicji informacji. Jedyny jej związek z informacją wynika z przekonania, że gdy znamy rozkład prawdopodobieństwa wyboru elementów pewnego zbioru (nazywanego w oryginalnym kontekście alfabetem), możemy określić miarę czegoś, co może być interpretowane jako informacja „wyprodukowana” (określenie Shannona) przez taki wybór. W przypadku entropii na gruncie teorii informacji mamy do czynienia jedynie z właściwościami ilościowymi.

Matematyczna teoria komunikacji (MTC) nie jest jedyną teorią, która szacuje liczbowo ilość informacji. Opierając się na wynikach osiągniętych w informatyce, niektórzy zaproponowali inne podejście do problemu ilości informacji. Ich badania są podstawą algorytmicznej teorii informacji. Kluczową ideą tej koncepcji jest założenie, że teoria, która wyjaśnia zjawisko X , jest swego rodzaju programem komputerowym, który oblicza X . Zatem taki program musi być w pewnym sensie mniejszy niż samo zjawisko, które wyjaśnia. Koncepcja informacji opiera się na takim stwierdzeniu i wynikach osiągniętych przy badaniu złożoności w informatyce.

Jednak informacja to nie jest tylko “coś w jakiejś ilości” tylko także “coś o czymś”.

W cybernetyce obok ilościowej charakterystyki informacji nawiązującej do zmniejszenia niepewności co do przewidywanych stanów układu pojawia się także próba jakościowego opisu informacji. Według Mariana Mazura informacja jest to transformacja poprzeczna komunikatów w torze sterowniczym, tj. zmiana pewnego stanu wyróżnionego w pewnym punkcie toru sterowniczego (systemu przekazywania informacji) na inny stan wyróżniony w tym samym punkcie; przyporządkowanie sobie tych

³ Claude Shannon, *Mathematical Theory of Communication*, 1948 r.

stanów. Czyli informacja zostaje określona formalnie jako relacja między tymi stanami. ⁴ Podkreślony zostaje w definicji informacji związek między stanami, element relacyjny między “czymś a czymś”

Podjęcie ilościowe Shannona czy czysto formalne, jakościowe Mazura nie rozstrzygały jednak kwestii semantycznej, dotyczącej znaczenia informacji, czyli w rozumieniu informacji jako “czegoś o czymś” tego drugiego “czymś”. Wspólną cechą teorii informacji semantycznej było zdefiniowanie semantycznej zawartości komunikatu w ramach pojęć teorii prawdopodobieństwa. Zarówno w teoriach Yehoushue Bara-Hillela i Rudolfa Carnapa czy Freda Dretskego oraz rozwijającej ją teorii Luciano Floridiego kluczowa w definicji pojęcia informacji semantycznej jest ocena wartości poznawczej, związanej z informacją, która jest mierzona odwrotnie proporcjonalnie do prawdopodobieństwa wystąpienia określonego stanu rzeczy. Zgodnie z tą zasadą, wartość informacyjna jest tym większa, im mniejsze ustalone prawdopodobieństwo wystąpienia określonego stanu rzeczy ze zbioru możliwości.

W teorii Bar-Hillela i Carnapa wartość ta jest ściśle związana z komunikatem językowym i jego miejscem w ramach konkretnego, zamkniętego systemu językowego. Jeśli informacja jest “czymś” o drugim “czymś”, to jest informacją semantyczna zdania A definiowana będzie jako to drugie “czymś” czyli zbiór światów możliwych wykluczonych przez A. Niektórzy podnoszą, że teoria Bar-Hillela i Carnapa nie jest wcale bardziej “semantyczna” od teorii ilościowych czy teorii Mariana Mazura. Zamiast akcentować ilość informacji, czy syntaktyczny/formalny związek między jakimiś stanami w pojęciu informacji opisują syntaktyczny związek z jakimś formalnie określonym zbiorem światów możliwych wykluczonych przez A.

W koncepcji Dretskego i jego następców wartość poznawcza modyfikowana jest natomiast przez subiektywne właściwości odbiorcy i chociaż nośnik informacji nie musi mieć charakteru językowego to samo pojęcie informacji zawężone zostaje do zdań prawdziwych. Treść informacyjną sygnału definiuje się następująco: Dla podmiotu x posiadającego wiedzę bazową K (o możliwościach istniejących u źródła informacji) sygnał r niesie informację, że s jest F w tw prawdopodobieństwo warunkowe tego, że s jest F pod warunkiem r (i K) jest równe 1 (ale pod warunkiem samego K jest mniejsze od 1). ⁵ Wypowiedzi fałszywe nie są według Dretskego informacjami.

Obok semantycznych teorii informacji odwołujących się do jakościowej reprezentacji powstawały także podobne teorie o charakterze pragmatycznym, odwołujące się do teorii decyzji. Informacja występuje tam jako bodziec dla określonego decyzji (“coś” jest bodźcem “czegoś”).

Niestety programowe, symboliczne ograniczenia tych klasycznych koncepcji (określające wymagania co do relacji między dwoma “czymś” lub charakteru określonego “czymś”), oraz ściśle związanych z klasycznym sposobem przetwarzania danych, sprawiają, że mają one bardzo wąskie zastosowanie.

W związku z niesamowitymi osiągnięciami i rozwojem nauk kognitywnych i sztucznej inteligencji (sztuczne sieci neuronowe) coraz częściej patrzy się na informację w kontekście przetwarzania przez takie sztuczne systemy. Informacja staje się po prostu “czymś” w relacji z “innym czymś”, jak relacje pomiędzy elementami neuropodobnymi. Na tak rudymenarnym poziomie dość jałowa staje się także rozmowa o semantycznym czy pragmatycznym charakterze informacji, semantyka czy pragmatyka staje się po prostu owocem syntaktyki, samych związków pomiędzy kolejnymi elementami sieci. Można powiedzieć, że **metafizyka współczesnego pojęcia informacji, kładzie bardziej nacisk na związek, relację między jakimiś elementami, niż na te elementy z osobna.**

Podobne intuicje daje nam współczesna fizyka kwantowa. Informacja kwantowa zakodowana jest w nielokalnych związkach pomiędzy poszczególnymi elementami układu, które nie mają odpowiednika w klasycznej teorii informacji. Mówimy wtedy, że układ jest w stanie splątany.

⁴ Marian Mazur, *Cybernetyczna teoria układów samodzielnych*, 1966, s. 35–37., Marian Mazur, *Jakościowa teoria informacji*, 1970, s. 70–71., Marian Mazur, *Cybernetyka i charakter*, 1976, s. 120–121

⁵ Dretske F., (1981), *Knowledge and the Flow of Information*, Cambridge, MIT Press, str. 65

2. Umysł i świadomość, czyli kim jesteśmy, jak postrzegamy i jak czujemy?

2.1. Umysł inteligentny, wiedza, myśl, myślenie

Dusza

Zasadniczo do połowy XX w. filozofia umysłu miała transcendentalny, fundamentalny i fenomenalny charakter i zarówno na gruncie religii i filozofii nawiązywała do pojęcia duszy.

Charakterystycznymi poglądami filozoficznymi na problem relacji duszy i ciała są poglądy Kartezjusza. Choć już Kartezjusz przyjmuje, „iż ciało jest zaledwie posągami i maszyną” to jednak znajduje także w szczytnym miejscu dla duszy. Ten szczególnie delikatny rodzaj subtelnej materii wprowadzany jest w drgania przez drgania duszy. Płyny cielesne przejmują te drgania, które poruszają następnie, organy, gruczoły i mięśnie. Z kolei Baruch Spinoza głosił, że istnieje tylko jedna substancja (Bóg), umysł i ciało są dwoma przejawami (własnościami) tej substancji. Gottfried Leibniz głosił z kolei, że monady obdarzone są dwoma atrybutami: rozciągłością i mentalnością. Synchronizacja działania monad możliwa jest dzięki harmonii przedustawnej. Wszystko zbudowane jest z monad, dlatego też wszystko obdarzone jest proto-mentalnością. Fenomenalny lub transcendentalny charakter umysłu był także przewodnim tematem późniejszej filozofii (np. fenomenologia) czy psychologii (np. fenomenologia, psychologia analityczna Junga).

Obliczenie

Rozwój cybernetyki, informatyki, sztucznej inteligencji, kognitywistyki (neurobiologii, psychologii poznawczej) oraz technologii komputerowej w połowie XX w. doprowadził do powstania wielu nowych dyscyplin i następnie do niesłychanie szybkiego rozwoju naszej wiedzy o ludzkim umyśle. Szczególnie w zakresie jego inteligencji; reprezentacji wiedzy oraz przetwarzania przez niego informacji.

W 1960 roku Hilary Putnam opublikował pracę *Minds and Machines*, broniąc w niej funkcjonalizmu maszynowego, koncepcji filozoficznej, która stała się dominującym nurtem w filozofii umysłu. Standardowy funkcjonalizm obliczeniowy (maszynowy) uwzględnia rolę obliczeniowe utożsamiane z określonymi rolami przyczynowymi w systemie poznawczym, a bodźce i zachowanie określa mianem „wejścia” i „wyjścia”. Z czasem z funkcjonalizmu obliczeniowego wyewoluował koneksjonizm, który nadal upatrywał fenomen inteligentnego umysłu w charakterze funkcji układu obliczeniowego, ale nie realizowanego w dowolny sposób a w sposób podobny do tego jak działa ludzki mózg. W obydwu przypadkach sposób rozumowania/funkcjonowania umysłu ma charakter mechaniczny/fizyczny (z powodu swojego obliczeniowego/fizycznego charakteru), natomiast w przypadku koneksjonizmu istotny staje się dodatkowo struktura połączeń pomiędzy neuropodowymi elementami sieci, czyli konkretne fizyczne jej ucieleśnienie.

Choć sam argument przeciwko tezie o obliczeniowym charakterze, czy o mechaniczności rozumowania ludzkiego przypisuje się Lucasowi to wiele osób, zapoznając się z wynikami Gödla, miało wcześniej poczucie, że dowiedzione jest ograniczenie dotyczące szeroko rozumianych maszyn, czyli komputerów, robotów i ich sieci (takie uwagi wypowiadali Kurt Gödel, Alan Turing czy Alonzo Church). Jednak dopiero w 1961 r., gdy postęp technologiczny konsekwentnie otwierał nadzieje na swój nieograniczony rozwój oraz kiedy pragmatycznie zorientowane dziedziny matematyki jako wspomniana matematyczna teoria informacji, cybernetyka, kognitywistyka, lingwistyka, psychologia poznawcza czy sztuczna inteligencja przybliżały coraz bardziej wizję „człowieka cyfrowego” niejako „w obronie wyjątkowości człowieka” ten filozoficzny argument został po raz pierwszy opisany w miarę

szczegółowo. Dlatego, pomimo tego, że sam tekst zawierający ten argument jest niezbyt dobry i tego, że autor nie należy bynajmniej do najwybitniejszych znawców twierdzenia Gödla mówi się zwykle o argumentacie Lucasa.

Argument Lucasa da się streścić w jednym zdaniu: Żadna maszyna nie może być równoważna umysłowi, bo umysł rozpoznaje prawdziwość zdania Gödla dla niej, a sama maszyna – na mocy twierdzenia Gödla – nie może, chyba że jest sprzeczna, ale wtedy na pewno nie jest równoważna umysłowi ludzkiemu. To wystarcza, by obalić tezę mechanicyzmu. Jest to wywód nad wyraz atrakcyjny. Wydaje się, że ściśle, niemal matematycznie, udowodniliśmy coś głębokiego na temat natury człowieka. Nic dziwnego, iż Lucas kończy swój tekst stwierdzeniem, że „jeśli ten dowód fałszywości mechanicyzmu jest słuszny, to ma to olbrzymie konsekwencje dla całej filozofii.” Nie będzie już presji wynikającej z rozwoju wiedzy przyrodniczej, by „negować wolność w imię nauki.” Bo „Żadne badania naukowe nie mogą wyczerpać nieskończonej różnorodności umysłu ludzkiego”.⁶

Argument Lucasa można ostatecznie sprowadzić do próby wykazania różnicy pomiędzy maszyną (systemem deterministycznym i symbolicznym) a człowiekiem poprzez wskazanie, że ta pierwsza jest ograniczona w swoim poznaniu a człowiek nie jest w nim ograniczony.

Dość bulwersującym faktem jest to, że w zasadzie argument ten zakłada to, co w rzeczy samej powinno być dowiedzione. Po pierwsze, można mieć wątpliwości skąd bierze się założenie, że aktualne poznanie czy doświadczenie człowieka nie jest w żaden sposób ograniczone (dowodzona wyższość człowieka została założona). Po drugie jeśli już uznamy, że poznanie człowieka uwikłane jest w potencjalnie nieskończone procesy ewolucyjne, albo związane jest z nieprzewidywalnymi efektami kwantowymi to dlaczego nie spojrzeć łaskawiej na samą maszynę, która także może lub mogłaby podlegać prawom ewolucji tak samo jak człowiek, czy “korzystać” z efektów kwantowych?

Należy uczciwie stwierdzić, że właściwie wśród logików matematycznych dominuje pogląd o nietrafności argumentu Lucasa, wynikającego rzekomo z prac Kurta Gödla, Alana Turinga czy Alonzo Churcha. Taka sytuacja ma miejsce począwszy już od 1960 r., kiedy to jeszcze przed publikacją Lucasa Hilary Putnam przedstawił swoją krytyczną wzmiankę uznaną Boolosa za „klasyczną”. Podobne poglądy znajdziemy w opinii Quine’a, analizach Benacerrafa i Wanga, czy w recenzji Putnama z wersji argumentu rozwiniętej przez Penrose’a.⁷

Ucieśnienie

Chociaż argumenty pokroju Lucasa zostały odparte, to początkowy optymizm panujący wśród

badaczy sztucznej inteligencji osłabł. Powodem były coraz większe trudności techniczne, które stawały na drodze do zrekonstruowania zdolności poznawczych człowieka za pomocą symboli.

Pierwsze architektury poznawcze (symulujące aspekty ludzkiego poznania) były to klasyczne systemy regułowe. Oparte były one o paradygmat przekształcania symboli wejściowych na symbole wyjściowe,

jednak symbole te związane były przede wszystkim ze zdaniami języka naturalnego (stąd ściśle związki

kognitywistyki z lingwistyką) lub formalnego a sama informacja była przetwarzana szeregowo. W latach 1955 – 1956 Herbert Simon i Allen Newell opracowali program (Logic Theorist) do dowodzenia twierdzeń logiki, który w przestrzeni możliwych przekształceń składniowych szuka niebanalnych dowodów. Uchodzi on za pierwszy program z zakresu sztucznej inteligencji. Program wyszukał twierdzenia 38 z pierwszych 52 twierdzeń Principia Mathematica Bertranda Russella. W programie do procesu skutecznego wyszukiwania dowodów wykorzystano także heurystyki, które ograniczały przestrzeń potencjalnych poszukiwań. Kolejnymi klasycznymi architektuрами z zakresu sztucznej inteligencji były General Problem Solver (1957), opracowywany pod kierunkiem Newella SOAR czy ACT Johna R. Andersona. Pomimo niezaprzeczalnego wkładu w rozumienie naszych procesów

⁶ John R. Lucas:// Umysły, Maszyny I Gödel, Hybris Nr 8 (2009) Issn: 1689-4286

⁷S. Krajewski 2003: *Twierdzenie Gödla i jego interpretacje filozoficzne: od mechanicyzmu do postmodernizmu* (Wyd. Instytutu Filozofii i Socjologii PAN, ISBN 83-7388-017-8)

poznawczych architektury ten nie spełniły wygórowanych oczekiwań najgorętszych ich entuzjastów. Raz po raz mnożyły się problemy związane z osiągnięciem celu ostatecznego, czyli budowy sztucznej inteligencji równej człowiekowi (General Artificial Intelligence). Coraz częściej podnoszono, że mieszcząca się w tym pierwszym, klasycznym paradygmacie kognitywistyki (Good Old Fashioned Artificial Intelligence) wizja natury umysłu ludzkiego jest niekompletna. Aby inteligencja mogła istnieć potrzeba z pewnością więcej niż pojedynczego mózgu. Inteligencja póki co powstaje w interakcji żywych organizmów z otoczeniem. Dlatego Hubert Dreyfus wprowadził do dyskusji dwóch filozofów, którzy o tej porze wydawali się być bardzo kognitywistycznej perspektywie oddaleni. Tymi filozofami byli Martin Heidegger i Maurycy Merleau-Ponty (fenomenologia). Głównym argumentem Dreyfusa przeciwko możliwości sztucznego myślenia jest esencjonalna nieformalizowalność całości ludzkiego doświadczenia. Ponadto w 1969 r. w logistycznym nurcie badań nad sztuczną inteligencją pojawił się konkretny problem (problem ramy)⁸. W tym nurcie tworzy się systemy dedukcyjne, które mają dostarczać narzędzi reprezentacji świata, zaś dzięki adekwatnej reprezentacji maszyna ma podejmować inteligentne decyzje. Reprezentacja wiedzy ma charakter zdaniowy i opiera się najczęściej na jakiejś wersji logiki pierwszego rzędu. Fakty w świecie wyrażane są przez tezy budowane przy użyciu zdefiniowanych aksjomatycznie predykatów. Na podstawie zbioru aksjomatów, opisujących podstawowe regularności w świecie (np. prawa mechaniki) i typy działań podmiotu inteligentnego (np. przesuwanie przedmiotów), wprowadzonych do bazy danych faktów i reguł działania "sztuczny inteligent" ma planować racjonalne i inteligentne decyzje, a następnie na ich podstawie działać. Okazuje się tylko, że tej wiedzy trzeba do tej bazy zapisać bardzo wiele (np. jeśli prasujemy koszulę w kratę o jej rękaw nie staje się dłuższy). Mianowicie, żeby z aktywnej wywołanej sytuacji mogły zostać wyprowadzone wnioski, następujące sytuacje muszą być szczegółowo opisane, w szczególności musi zostać powiedziane, co nie zachodzi w danej sytuacji. W normalnych sytuacjach wie się dużo na temat danej sytuacji, co jest w niej normalne, zwyczajne czy takie samo. Chociaż powszechnie przyjmuje się, że problem ramy nie jest logicznym czy filozoficznym argumentem przeciwko funkcjonalizmowi obliczeniowemu z tych samych przyczyn z jakich nie jest takim argumentem twierdzenie Lucasa (wystarczy wyobrazić sobie Eskimosa w Nowym Jorku lub nowojorczyka na biegunie północnym), to same trudności techniczne związane z efektywnością ówczesnej sztucznej inteligencji pozostawały faktem.

Aby pozbyć się problemu ramy rozpoczęto tworzenie logik, które były bardziej zbliżone do rozumowań zdroworozsądkowych⁹, z drugiej szybko rozwijające się badania neurofizjologiczne doprowadziły do powstania nowego podejścia do przetwarzania danych opartego na równoległym przetwarzaniu danych. U podstaw wczesnego rozwoju sztucznej inteligencji leżała między innymi praca z roku 1943, w której neurofizjolog Warren McCulloch oraz matematyk Walter Pitts wskazywali na podobieństwa między funkcjonowaniem neuronu a prostym, elektrycznie realizowanym obwodem logicznym. McCulloch miał przeświadczenie, że ludzki mózg jest komputerem, który dodaje dane wejściowe oraz porównuje je z daną wartością progową. Opisana w owej pracy struktura matematyczna może być postrzegana jako bezpośredni prekursor dzisiejszych sztucznych sieci neuronowych. W 1949 roku Donald Olding Hebb rozwinął strategię treningu dla naturalnych neuronów, która stała się znana pod nazwą „Reguła uczenia się Hebba”. Hebb przypuszczał, że uczenie się i ćwiczenie pamięci skorelowane jest ze zmianami siły połączeń pomiędzy pojedynczymi synapsami. Zgodnie z tą regułą, jednoczesna aktywacja neuronów prowadzi do wzmocnienia połączeń synaptycznych między nimi ("cell that fire together, wire together"). Kolejnym ważnym etapem prac nad sztucznymi sieciami neuronowymi stanowiły prace Franka Rosenblatta na temat perceptronu. Jemu i jego współpracownikom udało się konstrukcja pierwszego neurokomputera dzięki zastosowaniu w laboratoriach MIT implementacji założeń neuronalnych do odpowiednich „hardwarowych” komponentów. Chociaż badania Rosenblatta dotyczące sieci neuronowych zostały ostro skrytykowane przez Marviną Minskyego i Simona Paperta z powodu

⁸ John McCarthy and Patrick J. Hayes SOME PHILOSOPHICAL PROBLEMS FROM THE STANDPOINT OF ARTIFICIAL INTELLIGENCE Computer Science Department Stanford University Stanford, CA 94305 jmc@cs.stanford.edu <http://www-formal.stanford.edu/jmc/> 1969

⁹ Już sam McCarthy w ramach rachunku sytuacji opracował pewną wersję logiki niemonotonicznej, będącej modelem rozumowań zdroworozsądkowych

plaskości tych sieci, to szybko okazało się, że krytycy nie docenili doniosłości nowego podejścia.¹⁰ Dopiero w latach 80 tych XX wieku opracowano ostatecznie skuteczną metodę uczenia się wielowarstwowych takich jak perceptron (metoda propagacji wstecznej).¹¹ Powrót do badań nad sieciami neuropodobnymi nastąpił min. za sprawą pracy Johna Hopfielda, który stworzył sieć ze sprzężeniem zwrotnym rozwiązującą dosyć trudne zadania (np. problem komiwojażera o wysokiej złożoności obliczeniowej). Stworzono wiele innych modeli sieci neuropodobnych o różnych zastosowaniach, np. sieci wykorzystujące logikę rozmytą oraz sieci modelujące pamięć asocjacyjną. Opracowano też różnego rodzaju algorytmy ich uczenia.

Następujący wraz z rozwojem nauki rozwój technologii powodował, że orędownicy nowego podejścia w kognitywistyce opartego na sieciach neuropodobnych zaczęli znów wierzyć, że są ostatecznie są one w stanie dorównać w zakresie inteligencji ludzkim neuronom. Funkcjonalizm obliczeniowy zaczął być wypierany z pozycji dominujących przez koneksjonizm, czyli funkcjonalizm koneksjonistyczny. Wraz z kolejnymi spektakularnymi sukcesami sztucznej inteligencji i uczenia maszynowego (w szczególności *deep learning*) takimi jak wygrana w szachy z Garym Kasprosem w 1997 r., wygrana w GO z Lee Sedolem w 2016 r. czy wreszcie prezentacji ChatGPT w 2022 r. coraz częściej każdemu z nas towarzyszy myśl, że nic nie stoi na przeszkodzie aby wygenerować równą człowiekowi inteligencję w sposób niekoniecznie taki sam jak zrobiła to sama natura.

Jednak dominujące we współczesnej nauce o umyśle przekonanie, że umysł co do zasady jest zdeterminowany przez swoją funkcję obliczeniową nie pozostawiało w spokoju wielu myślicieli. Jeszcze w 1980r. John Searl, przedstawiciel dualizmu naturalistycznego, przedstawił w swojej pracy zatytułowanej „Minds, Brains, and Programs” argument chińskiego pokoju. Istotą argumentu jest pokazanie, że skuteczne symulowanie rozumu przez maszynę nie jest tożsame z jego posiadaniem, ponieważ wykonywanie określonych zadań nie musi wiązać się z ich rozumieniem przez wykonawcę. Dość powszechnie uważa się, że Searl upatrujący świadomości i zdarzeń umysłowych jako realnie istniejących biologicznych własności mózgu (i jako realnie istniejących mogących działać przyczynowo) wydawał się jednak nie rozumieć samego fenomenu złożoności i możliwych konsekwencji wynikających z funkcjonowania złożonych układów neuropodobnych. Podobnie jak to było w przypadku informacji, **związki pomiędzy elementami okazują się mieć fundamentalne znaczenie dla rozumienia i samej semantyki.** Już dzisiaj, rozmawiając z chatem GPT wielu z nas ma po prostu poczucie, że rozmawia z systemem, który w coraz większym stopniu “rozumie” o czym rozmawia. Wobec sukcesów, która odnosiła sztuczna inteligencja w zakresie myślenia, fenomenalne aspekty Umysłu coraz częściej dyskutowane były w ramach problematyki innych fenomenów ludzkiego *psyche*, takich jak emocje i przede wszystkim wszystkim wrażenia i wreszcie **Świadomość.**

2.2. Emocje

Cielesna dusza, ludzki ruch zamiast duch

Problem emocji w filozofii nigdy nie odgrywał pierwszorzędnej roli. Poniekąd uprzywilejowanym przedmiotem filozofii (a zarazem narzędziem) był rozum, jego czynności i wytwory - „wiedza raczej aniżeli działanie, przekonanie aniżeli emocje, intelekt aniżeli wola były centralnymi tematami badań filozoficznych”.¹² Z emocjami związane było greckie słowo πάθος (pathos), pochodzące od czasownika pati – ‘doznać’. Podobnie jak łacińskie ‘Passiō’, które zwykle tłumaczone jest jako ‘uczucie’, ‘namiętność’, ale też ‘cierpienie’ lub ‘męka’. Z emocjami związane jest także inne łacińskie słowo: ‘e movere’ oznaczające ‘umysł w ruchu’. Przez wieki emocje utożsamiane z jednej strony jako element irracjonalny, zmienne doznanie duszy i przeciwstawiane rozumowi a z drugiej strony jako element mocno związany z ciałem. Dopiero narodziny psychologii empirycznej w XXw.

¹⁰Minsky i Papet wykazali, że perceptron jednowarstwowy nie może zrealizować operacji alternatywy rozłącznej. Chociaż już wtedy wiadomo było, że taką operację jest w stanie zrealizować perceptron dwuwarstwowy; podejście określili oni jako jałowe i na długi czas zatrzymali rozwój nad sieciami neuronowymi.

¹¹Aktualnie wszyscy przyznają, że siła sieci zależy od tego w jaki sposób połączone są węzły sieci

¹² Kenny A. (1963), Action, Emotion and Will, London: Routledge and Kegan Paul.,

spowodowały, że ten status emocji jako „ucieleśnionej duszy”, został zracjonalizowany ze szczególnym uwzględnieniem ucieleśnienia.

Standardowa Charakterystyka Emocji zakłada w pewnym uproszczeniu, że typowa emocja jest: (i) bardzo szybkim, (ii) skorelowanym ze zmianą stanu ciała, (iii) mimowolnym zachowaniem się, (iv) będącym reakcją, (v) na zidentyfikowany percepcyjnie, (vi) i zwaloryzowany (pozytywnie lub negatywnie), (vii) obiekt (stan rzeczy) występujący w otoczeniu. Emocjom towarzyszy aspekt fenomenalny, związany z odczuwaniem „stanu afektywnego (felt affective state).¹³

Klasyczne teorie emocji ukierunkowane były przede wszystkim na pobudzenia fizjologiczne połączone z fenomenalnym odczuwaniem niż na abstrakcyjne poznanie lub ocenę. Spór między przedstawicielami teorii klasycznych sprowadzał się przede wszystkim do ustalenia wagi lub kolejności występowania w procesie uruchamiania emocji poszczególnych elementów ją definiujących (Teoria Jamesa-Langego, teoria Cannona-Philipa Barda, Teoria aktywacyjna, dwuczynnikowa teoria Stanleya Schachtera i Jerome’a Singera).

Obliczenia

Teorie poznawcze emocji stanowią bogaty wachlarz nie tylko w wymiarze teorii psychologicznych, ale także filozoficznych (Errol Bedford, William Lyons, Martha Nussbaum, Robert Solomon i wielu innych). W przeciwieństwie do wcześniejszych koncepcji ukierunkowanych na pobudzenia fizjologiczne i odczucia z nimi związane, w teoriach poznawczych emocje utożsamiane są w nich zwykle z myślami, przekonaniem lub sądami albo z ich podstawowymi komponentami. Zwolenników różnych teorii poznawczych łączą trzy hipotezy: 1) konceptualizacja; 2) bezcielesność oraz 3) ocena.¹⁴ Tym samym emocja staje się jakby strukturą poznawczą pojęciową jak myśl (opisana w poprzednim rozdziale) a z drugiej dotyczy osądu (pozytywnego lub negatywnego) pod kątem zaspokajania potrzeb. Szczególnie ważną odmianą poznawczych koncepcji jest, powstałe w ramach filozofii analitycznej, rozwiązanie propozycjonalne. Za jego twórcę uważa się Bertranda Russella. Postawy propozycjonalne (inaczej „nastawienia sądownicze”) to stany intencjonalne podmiotu, proste myśli, m.in. przekonania, pragnienia, emocje itd., wyrażane przez podmiot w języku. „P twierdzi, że x”, „P obawia się, że y”, „P kocha Z” itd. W każdej z takich postaw można wyróżnić czasownik, klauzulę „że” oraz zdanie o pewnej treści x. Stosując to rozwiązanie do emocji, przyjmuje się, że postawy emocjonalne są przede wszystkim świadomie i intencjonalnie nakierowane na pewien obiekt. W koncepcji tej emocje są w swojej esencji traktowane jako element racjonalnego myślenia (poznania), którego podstawowymi składnikami, powtórzmy, są pojęcia i przekonania.¹⁵ Dlatego też, nie ma większych przeszkód, aby emocje mogły być także dostępne maszynom w procesie przetwarzania informacji. Teoriom poznawczym zarzuca się, że doznający emocji nie musi zawsze dysponować poznaniem jakiegoś przedmiotu lub żywić określonych przekonań, by zareagować emocjonalnie. Po drugie może on doznać emocji zanim coś rozpozna.

Poszczególne teorie poznawcze (teoria Richarda Lazarusa, teoria Nico-Frijdy, teoria komunikacyjna Keitha Oatleya i Philipa Johnsona-Lairda) różnie rozkładają wagę co do konkretnych elementów oceny lub procesu poznawczego.

Ucieleśnienie

Koncepcja Jessego Prinza jest próbą unifikacji dotychczasowych podejść, u podstaw której leży jednak przekonanie o potrzebie utrzymania standardowej charakterystyki emocji. Z jednej strony koncepcja Prinza sympatyzuje z fizjologiczną teorią Jamesa-Langego, z drugiej podkreśla także

¹³ D. Wiener, Wprowadzenie do Emocje: domysły, pomysły i fakty w formy aktywności umysłu, w emocje, percepcja, świadomość cz. 1., str. 1-11

¹⁴ J. Prinz, Gut Reactions. cyt., s. 21-26

¹⁵W: Dąbrowski A. CZYM SĄ EMOCJE? PREZENTACJA WIELOSKŁADNIKOWEJ TEORII EMOCJI, „Analiza i egzystencja” 27 (2014) ISSN 1734-9923

poznawczy charakter emocji. Prinz przyjmuje, że emocje powstają w ciele – są cielesnymi zmianami. Podstawową funkcją emocji jest rejestracja (percepcja) wzbudzonych stanów ciała i ich ocena: emocje to ucieleśnione oceny. Ocenić x to ująć x w formie reprezentacji. A zatem emocje są nie tylko ocenami, ale też reprezentacjami: „Mówiąc, że emocje są spostrzeżeniami zmian cielesnych, mam na myśli jedynie to, że są takie stany w naszych układach sensomotorycznych, które rejestrują zmiany w naszych ciałach. [...] Pomocne jednak będzie odróżnienie ‘rejestrowania’ od ‘reprezentacji’. [...] Stan umysłu rejestruje to, co niezawodnie powoduje jego aktywację. Emocje najwyraźniej ‘rejestrują’ zmiany w ciele, lecz istnieje jeszcze kolejne pytanie dotyczące tego, co te stany reprezentują”¹⁶

Koncepcję Prinza rozwija następnie w kierunku dalszego ucieleśnienia Jaak Panksepp. Koncepcja Pankseppa do dziś pozostaje jednym z najistotniejszych prób ujęcia problematyki emocjonalnej na podstawie neurobiologicznej, inspirowanej psychologią, kognitywistyką i filozofią. Panksepp jest przekonany o autonomii sfery poznawczej i emocjonalnej. Uważa on, że emocje tworzą swoisty neurobiologiczny rodzaj naturalny w mózgu ssaków.¹⁷ Panksepp zaproponował także rozróżnienie następujących, podstawowych mózgowych systemów emocjonalnych: POSZUKIWANIA, WŚCIEKŁOŚCI, STRACHU, POŻĄDANIA, OPIEKI, PANIKI/SEPARACJI oraz ZABAWY (Panksepp rozmyślnie używa dużej czcionki).

W przypadku poznawczego aspektu emocji, podobnie jak w przypadku procesu myślenia, wydaje się, że budowa odpowiednich neuropodobnych sieci (**konkretne ucieleśnienie systemu obliczeniowego oparte na relacjach pomiędzy elementami neuropodobnymi**) wystarcza do jego skutecznego odtworzenia. Jednak aktualne pozostaje pytanie o treść towarzyszących doznań (także tych afektywnych) i towarzyszącej temu procesowi określonej fizjologii. W przypadku emocji, jako tak ściśle związanych z ciałem, pytanie to staje się jeszcze bardziej kluczowe.

Czy zupełnie inne reakcje, inne ich odczuwanie (jeśli na moment założymy, że takie odczuwanie jest w ogóle możliwe) i inne odczuwanie stanów afektywnych np. w ramach sztucznych systemów pozwoliłoby nam w ogóle dostrzec jakąkolwiek emocjonalność? Czy definicje emocjonalności ograniczamy tylko do podobnej naszej? Ucieleśnionej dokładnie tak samo jak nasza? Podobnie? Jak podobnie? Dlatego też w jeszcze większym stopniu niż w przypadku myślenia czy inteligencji, poza wspólnym aspektem poznawczym (gdzie sprawa ogranicza się do wykonania określonych obliczeń) szczególnie istotna staje się refleksja nad tym jak istotny dla nas jest element jakościowy (i związane z nim ucieleśnienie). Pytania te otwierają w naturalny sposób drogę do omówienia wrażeń i świadomości, jako pojęć najbardziej chyba uwikłanych w fenomenologiczne lub ucieleśnione ujęcia (ze względu na swoją lub ich przedmiotu nieuchwytność).

2.3. Wrażenia, Świadomość

Dusza

Jak wspomniane zostało już wcześniej zasadniczo do połowy XX w. filozofia umysłu a także filozofia świadomości miały transcendentalny, fundamentalny i fenomenalny charakter i zarówno na gruncie religii i filozofii nawiązywały do pojęcia duszy. Jednak kiedy w połowie XX w. na skutek rozwoju kognitywistyki i sztucznej inteligencji problem myślenia i bardziej z nim związanego Umysłu dominował w debacie naukowej, problem Świadomości jako problem *stricte* metafizyczny przestał na moment przykuwać uwagę jej uczestników.

Doświadczenie zmysłowe jest pierwotną lokalizacją świadomości. Niektórzy uważają, że bez

¹⁶ J. Prinz, Gut Reactions. cyt., s. 58

¹⁷ Panksepp J, Emocje jako twory naturalne w mózgu ssaków. W: M. Lewis, J.M. Haviland-Jones. Psychologia emocji (str. 185-210), Panksepp J. Affective Neuroscience: the Foundations of Human and Animal Emotion. New York, NY: Oxford University Press. 1998 str. 46

doświadczenia zmysłowego nie mielibyśmy pojęcia świadomości.¹⁸ Niezależnie od tego, czy jest to prawdą, doświadczanie zjawisk, wyglądu, dźwięku, smaku i czucia dominuje w naszym życiu mentalnym. Jak pisze Fred Dretske: Całkowicie usunięcie tego doświadczenia sprawiłoby, że stalibyśmy się kim? Zombie?¹⁹

Percepcja to jedna z najważniejszych i najlepiej poznanych form aktywności umysłu. Badania tej aktywności nie tylko dostarcza wiedzy o funkcjonowaniu wyspecjalizowanych systemów zbierania informacji o otoczeniu, ale także ukazuje działanie umysłu w jego najbardziej fundamentalnych związkach ze światem. Badanie procesów percepcyjnych jest jednym z filarów całej kognitywistyki. Celem percepcji jest zdobycie przez perceptor informacji o przedmiotach, stanach lub cechach szczególnie środowiska. Informacja ta ma zasadnicze znaczenie dla przetrwania perceptorów oraz pozwala reagowanie na zmieniające się warunki otoczenia.

Percepcja to złożony proces. Etapem wstępnym jest zebranie materiału do obróbki percepcyjnej. Następnie materiał zostaje poddany dalszemu, wieloetapowemu przetworzeniu. Na kolejnym etapie perceptor dokonuje wyboru danych, na których koncentruje swoją uwagę. Na dalszych etapach przetwarzania danych włącza się także dodatkowy składnik, mianowicie uprzednio zdobyta wiedza perceptorów. Problem udziału tej wiedzy stanowi wyzwanie dla współczesnych badaczy percepcji. To dodatkowe sprzężenie pomiędzy tym co dostarczane a co już znane zaburza obraz percepcji jako procesu jednokierunkowego. Można jednak przyjąć, że ostatecznym rezultatem tego procesu jest percept, stan umysłu, za pomocą którego umysł „odzwierciedla” określony fragment rzeczywistości. Jest więc jego reprezentacją. Przekonanie o informacyjnym charakterze jest powszechnie przyjmowane.

Oprócz problemu związanego z reprezentacją, najtrudniejszym zagadnieniem jest problem subiektywnych wrażeń lub ściślej rzecz biorąc problem specyficznych, subiektywnych jakości, jakie wiążą się z wrażeniami. Jest to stary problem przedstawiony w nowym ujęciu przez Davida Chalmersa²⁰ i omawiany na łamach „Journal of Consciousness Studies”. Tym samym problem Świadomości wkroczył na grunt nauki. Łacińskie słówko qualis oznacza własność w oderwaniu od rzeczy, posiadających daną własność, stąd mówi się potocznie o „problemie kwaliów”.

Kiedy wyobrazimy, że leży przed nami duża truskawka, skupiamy się na jej barwie. Rozpoznamy, że jest czerwona, a identyfikacji tej cechy towarzyszy zmysłowe doznanie czerwieni. Czujemy intuicyjne, że rozpoznanie i świadome doznanie tej barwy to dwa różne procesy. Ten pierwszy wynika z działania urządzenia pomiarowego służącego do rozpoznawania barw w postaci podukładu wzrokowego układu sensorycznego, poczynając od czopków na siatkówce oka, a kończąc na polu V4 w korze wzrokowej. O doznaniu można jedynie powiedzieć, że współwystępuje ono z efektem końcowym tych procesów, czyli rozpoznanie czerwonej barwy. To subiektywne doznanie związane jest z drugim, fenomenologicznym aspektem świadomości.

Istnienie tego subiektywnego aspekt wrażeń (ale i także emocji i nawet samego myślenia) dla wielu także stanowi poważny argument przeciwko możliwości ujmowania fenomenu świadomości człowieka tylko w kategoriach obliczeniowych, nawet jeśli ucieleśnionych w sieciach neuropodobnych.

Obliczanie

Dla zwolenników funkcjonalizmu obliczeniowego i koneksjonizmu wrażenia były albo problemem nie istotnym albo wrażenia także pozostawały w relacji funkcyjnej do samych obliczeń, realizowanych przede wszystkim w złożonych i skalowalnych sieciach neuropodobnych. Ucieleśnienie w neuropodonej formie staje się kluczowym warunkiem do stwierdzenia jakościowej tożsamości

¹⁸ Marcel A.J.E. BISIACH, Phenomenal experience and functionalism. W: Marcel, Bisiach 1988, Consciousness in Contemporary Science. Oxford: Clarendon Press, str. 121-1581

¹⁹ F. Dretske, Naturalizowanie umysłu, str. 19

²⁰ David Chalmers, (1996), *The Conscious Mind*, New York: Oxford University Press, ISBN 9780195117899 ,

między wrażeniami układu i wrażeniami człowieka (także doznań związanych z emocjami). Drugorzędną kwestią staje się zarówno kwestia złożoności i konkretnej budowy takiej sieci a także jej dalszego ucieleśnienia związanego z fizycznymi relacjami między organami a pośrednio także ze środowiskiem.

Ucieleśnienie

Jak wspomniałem już wcześniej (przy omawianiu problematyki Umysłu, wiedzy, myśli i myślenia a konkretnie ich ucieleśnienia) w latach 70tych i 80tych XX w. optymizm panujący wśród badaczy sztucznej inteligencji osłabł, ze względu na coraz większe trudności techniczne, które stawały na drodze do zrekonstruowania zdolności poznawczych człowieka za pomocą symboli. Według takich filozofów jak wspomniani Hubert Dreyfus czy John Searl głównym argumentem przeciwko możliwości sztucznego myślenia jest esencjonalna nieformalizowalność całości ludzkiego doświadczenia.

Kiedy pomimo pojawiających się problemów technicznych klasyczne kognitywistyczne ujęcie umysłu i świadomości ewoluowało ku ujęciu koneksjonistycznemu, jednak traktującego je jako maszynę obliczeniową (nawet jeśli o specyficznej budowie) to wciąż wielu próbowało ratować zawartą w naszych umysłach i świadomości tajemnicę ludzkiego niemaszynowego jestestwa. Niektórzy odwołując się do naukowego naturalizmu próbowało ignorować ten temat sprowadzając go wprost do dyskusji o właściwościach fizycznych (Paul M.Churchland). Niektórzy naturaliści odwoływali się do modeli wręcz dualistycznych i wprost fenomenalistycznych; gdzie ten subiektywny (fenomenalny) aspekt *a priori* istnieje obok elementu obiektywnego (John Searl, David Chalmers). Roger Penrose upatrywał z kolei istoty umysłu i świadomości w efektach kwantowych, czym rozbudził wyobraźnię wielu zwolenników najróżniejszych teorii fundamentalnych.²¹ Uczciwie należy jednak stwierdzić, że pogląd ten jest poglądem powszechnie krytykowanym przez większość przedstawicieli nauk kognitywnych, efekty kwantowe nie wydają się mieć większego znaczenia dla omawianych funkcji mózgu.

Równolegle rozwijał w ramach kognitywistyki nowy kierunek, zainspirowany Heideggerowskim „byciem w świecie”. Poszukiwanie istoty umysłu i świadomości w konkretnym ucieleśnieniu układu nerwowego, w szczególności jego interakcji z ciałem i środowiskiem stało się kamieniem milowym kognitywistyki ucieleśnionej. Rzecznikami tego podejścia zostali m.in. lingwista George Lakoff i filozof Mark Johnson, którzy sprzeciwili się ideom jednego z ojców-założycieli kognitywistyki, Noama Chomsky’ego. Ten ostatni twierdził, że w ludzkim umyśle istnieje specjalny obliczeniowy moduł, działający według abstrakcyjnych reguł i odpowiedzialny za przyswajanie języka oraz kompetencje językowe. Według Lakoffa i Johnsona: „nie istnieje ktoś taki jak człowiek obliczeniowy (...). Prawdziwi ludzie mają umysły ucieleśnione, a ich systemy pojęciowe powstają dzięki żywemu ciału, są przez nie ukształtowane i dzięki niemu posiadają znaczenie”. Innymi słowy, tożsamość idei ucieleśnienia wyznaczana była wówczas przez odrzucenie komputerowej wizji umysłu.

Jeśli sam funkcjonalizm ewoluował ze stanowiska czysto obliczeniowego do koneksjonistycznego uwzględniając postulat ucieleśnienia w zakresie określonej budowy elementów sieci neuropodobnej i samej sieci, to wciąż postulował także właściwość metafory komputerowej. Sami kognitywiści ucieleśnieni chcieli pójść znacznie dalej, uznać, że dla odtworzenia świadomości nie jest wystarczającym spełnienie neuropodobności w zakresie przetwarzania informacji. Istoty umysłu i świadomości chcą oni szukać w relacjach między systemem nerwowym a ciałem i środowiskiem (ucieleśnienie) albo w bardziej fundamentalnych prawach przyrody (efekty kwantowe). Jakkolwiek trudno odmówić racji, że ciało i środowisko (to co jest i jakimi się rządzi prawami) oddziałują na nasz umysł i świadomość metaforycznie stają się jego częścią, to trudno też zaprzeczyć, że gdyby można było przez chwilę wyjąć nasz mózg z ciała i utrzymując go przy życiu umieścić w próżni, zapewne jeszcze przez jakiś czas miałby myśli, emocje i wrażenia takie jakie właściwe są nam wszystkim. Dlatego też nie wydaje się być zasadne wyłączenie z zakresu myśli,

²¹ R. Penrose, Wielki Umysł Cesarza,

emocji i wrażeń mózgu, który miał zupełnie inne od naszego ciała lub/i inne środowisko. Nawet jak w takim przypadku taki umysł czy świadomość byłaby dla nas po prostu ponownie nierozpoznawalna.

Na sam koniec chciałbym jeszcze zaznaczyć, że zarówno funkcjonalizm (koneksjonizm) jak i naturalistyczna kognitywistyka ucieleśniona fenomen Umysłu czy Świadomości sprowadzają ostatecznie **do dynamicznej relacji pomiędzy elementami neuropodobnymi w pierwszym przypadku i relacji z ciałem czy otoczeniem w drugiej. W każdym przypadku prowadzi nas do wizji Świata, w którym nasz Umysł i Świadomość jawią się bardziej jako części Całości niż całkowicie osobną i autonomiczną rzecz.** Świat jest bowiem deterministyczny lub niedeterministyczny i nie w nim miejsca na jakąkolwiek inną autonomię.

3. Wartość, użyteczność, szczęście, dobro i piękno, czyli jak postępujemy lub powinniśmy postępować (lub jaki świat powinien być)?

3.1. Wolna wola, jaźń, instykt, intencja

Dusza, Wolna wola

Problem wolnej woli jest fundamentalnym dla rozstrzygnięcia o regułach ludzkiego postępowania. Niewątpliwie problem woli zależy także wprost od jeszcze bardziej fundamentalnego pytania dotyczącego przyczynowości w fizyce w ogóle.

Opis świata w ramach fizyki klasycznej, newtonowskiej z zasady jest opisem deterministycznym, nawet jeśli pojawiają się w niej pewne sugestie istotnej przypadkowości (sugestie takie wysuwał Maxwell w ramach badań nad ruchami cząsteczek w gazach). Mechanika kwantowa natomiast, przynajmniej na pierwszy rzut oka zdaje się wskazywać na coś zupełnie innego. W szczególności tzw. nierówności Bella mają jednoznacznie wykazywać, że u swoich fundamentów świat jest przypadkowy. Fundamentalny determinizm świata można jednak bronić w teoriach nielokalnych, czyli takich, w których istnieje możliwość przesyłania informacji szybciej niż światło. Przykładem takiej teorii jest teoria ukrytych parametrów Davida Bohma.

Niezależnie jednak czy świat jest u swoich podstaw całkowicie deterministyczny czy indeterministyczny trudno w ramach fizyki znaleźć uzasadnienie dla wolnej woli. Determinizm świata wprost wyklucza wolną wolę. Pomimo prób wiązania istnienia świadomości a także i wolnej woli z indeterministycznymi efektami kwantowymi (przez Rogera Penrose'a), dość powszechnie uważa się jednak, że indeterminizm (czyli całkowita przypadkowość) nie tylko nie ma nic wspólnego z wolną wolą, ale równie wprost jak determinizm ją wręcz wyklucza.

Niezależnie od refleksji nad wynikami nauk fizycznych, do czasu powstania współczesnych nauk o umyśle, z założenia istnienie wolnej woli było powszechnie akceptowane. Nawet Kartezjusz, który przyjmował „iż ciało jest zaledwie posągami i maszyną” jednak ostatecznie znajduje także (w szyszynce) miejsce dla duszy, a także jej aktywnego aspektu: wolnej woli. Ten szczególnie delikatny rodzaj subtelnej materii wprowadzany jest w drgania przez drgania duszy. Płyny cieleśne przejmują te drgania, które poruszają następnie, organy, gruczoły i mięśnie.

Obliczenia

Dopiero powstanie metafory komputerowej we współczesnej nauce o umyśle powszechnie zakwestionowało istnienie wolnej woli. Pomimo dość desperackich prób jej ugruntowania min. w efektach kwantowych przez wspomnianego wcześniej Rogera Penrose'a dość powszechnie uznaje się, że niemalże niemożliwym jest jej udowodnienie, poza powołaniem się wprost na jej poczucie (John Searl). Co więcej, pojawiają się kolejne poważne argumenty przeciwko jej istnieniu.

Jednym z najsłynniejszych argumentów przeciwko wolnej woli jest tzw. eksperyment Libeta.²²

W oryginalnym eksperymencie Benjamina Libeta brało udział pięciu badanych, których zadaniem było wykonywanie spontanicznych ruchów ręką. Moment ruchu miał zależeć tylko i wyłącznie od nich — Libet instruował uczestników, by poruszali nadgarstkiem w możliwie najbardziej spontaniczny sposób. W czasie trwania eksperymentu badani wpatrywali się w zegar składający się z tarczy oraz zielonego punktu, który obiegał ją w 2,5 sekundy. Ich dodatkowym zadaniem było zapamiętanie, w którym miejscu znajdował się zielony punkt, gdy zdecydowali się na wykonanie ruchu. Libet rejestrował aktywność mózgow badanych za pomocą EEG, chcąc w ten sposób wykazać, że świadome podjęcie decyzji pojawia się przed aktywnością neuronalną związaną z ruchem. Zgodnie z oczekiwaniami okazało się, że moment podjęcia świadomej decyzji poprzedzał ruch ręką o 200 milisekund, Libetowi jednak nie udało się potwierdzić istnienia wolnej woli. Symetryczny potencjał gotowości (RP), reprezentujący aktywność neuronalną obu półkul mózgowych (głównie w obszarach dodatkowego pola ruchowego, pierwotnej kory ruchowej i kory sensorycznej), pojawiał się średnio 550 milisekund przed ruchem ręką, a zatem 350 milisekund przed świadomym podjęciem decyzji. Świadoma decyzja, wymagana do istnienia wolnej woli, była więc poprzedzana nieświadomą aktywnością mózgu, z czego wynikał (przynajmniej na pierwszy rzut oka) prosty wniosek: wolna wola nie istnieje. W 2008 roku John-Dylan Haynes wraz z zespołem przeprowadził podobny eksperyment, jednak unowocześniony w praktycznie każdym aspekcie. Wyniki tego eksperymentu według wielu badaczy potwierdzały wnioski z eksperymentu Libeta.

Ucieśnienia: jaźń, instykt, intencja

Niezależnie od rozstrzygnięcia o realności wolnej woli i realnym jest podmiotowe poczucie istnienia i poczucie sprawstwa, które można określić aktywnym aspektem świadomości. Bez świadomości nie ma doświadczenia podmiotowości ani sprawstwa, nie pojawiają się cechy strukturalne jaźni, nie ma wymiaru fenomenalnego ani osobowego. Wydaje się, że mają rację autorzy, którzy twierdzą, że nie ma świadomości bez kogoś, kto jest świadomy. Tak zwana egotyczna orientacja jest cechą strukturalną wszelkich odmian oraz form rozwojowych świadomości (wyposażonych w element samoodniesienia). Powyższą tezę Galen Strawson określa mianem The Experience/Experencer Thesis i wyraża ją za pomocą trzech logicznie równoważnych twierdzeń: a) doświadczenie jest niemożliwe bez doświadczonego, b) doświadczenie jest z konieczności czymś doświadczeniem, c) podmiot doświadczenia istnieje tylko wówczas, gdy istnieje doświadczenie. W podobnym duchu pisze Antonio Damasio: Gdyby proces jaźni załamał się i ustał kompletnie, umysł straciłby orientację, zdolność do zebrania swych części. Myśli płynęłyby swym własnym torem, bez właściciela, który by się do nich przyznał. Nasza skuteczność w realnym świecie spadłaby do zera i zewnątrzni obserwatorzy uznaliby nas za straconych. Jak byśmy wyglądali? Cóż, wyglądalibyśmy na nieświadomych. Formy rozwojowe jaźni odpowiadają fazom rozwojowym świadomości: a) protoświadomości odpowiada protojaźń, b) świadomości rdzennej – jaźń rdzenna, zaś c) świadomości rozszerzonej – jaźń rozszerzona (autobiograficzna). Przedmiotem szeroko zakrojonych badań (empirycznych i teoretycznych) jest ustalenie, od jakiego momentu w rozwoju jednostek i gatunków możemy mówić o minimalnym podmiocie lub minimalnej jaźni. Wydaje się, że gdziekolwiek mamy do czynienia choćby z najbardziej elementarną formą świadomości, jest to świadomość przyporządkowana określone podmiotowi (Ja). Proste formy nierefleksyjnej samoświadomości oraz jaźni towarzyszą nam od pierwszych dni naszego istnienia. Ja cielesne, Ja ekologiczne, proto-Ja, Ja rdzenne i Ja utajone – to różne określenia pierwotnych, prostych, elementarnych form doświadczanej podmiotowości. Natomiast rozwój wyższych form samoświadomości i jaźni wymaga interakcji społecznych, doświadczania cudzych stanów psychicznych, języka, pamięci autobiograficznej itd. Jak pisze Damasio: „Jaźń zdecydowanie nie jest

²² Libet B. (1985). „Unconscious cerebral initiative and the role of conscious will in voluntary action”, *The Behavioral and Brain Sciences*, t. 8, s. 529-566.

rzeczą; jest dynamicznym procesem, utrzymywanym na dość stabilnym poziomie przez większą część okresu czuwania, ale podlegającym mniejszym lub większym zmianom przez cały ten czas. (...). Najprostszy poziom (protojaźń) jest dziełem tej części mózgu, która reprezentuje organizm, i składa się ze zbioru obrazów, które opisują stosunkowo stabilne aspekty ciała oraz wytwarzają spontaniczne odczucia żywego ciała (czucia pierwotne). Drugi etap jest wynikiem ustanowienia powiązania między organizmem (reprezentowanym przez protojaźń) a dowolną częścią mózgu, która reprezentuje obiekt poznawany. Efektem jest jaźń rdzenna. Na trzecim etapie (...) powstaje jaźń autobiograficzna, która definiowana jest w kategoriach wiedzy biograficznej, obejmującej zarówno przeszłość, jak i przewidywaną przyszłość. Wszystkie trzy etapy realizowane są w odrębnych, lecz skoordynowanych przestrzeniach roboczych mózgu”.

Jeśli pojawia się podmiot (opisując jego zachowanie) pojawia się pytanie o przyczyny, motywacje czy cel takiego zachowania. Najbardziej podstawowym mechanizmem odpowiedzialnym za kierunek ludzkiego zachowania jest instynkt. Instynkt (podnieta, popęd) to dyspozycja biopsychiczna^[1], która powoduje, że osobnik w nią wyposażony zachowuje się (lub pragnie się zachowywać) w określony sposób, a także "wrodzona zdolność wykonywania pewnych czynności stereotypowych, niewyuczonych, mniej lub bardziej skomplikowanych swoistych dla danego gatunku istotnych dla jego przetrwania". Pojęcie to w nauce o zachowaniu jest różnie definiowane, a w języku potocznym kojarzone jest z intuicją i zachowaniami, które nie są wynikiem racjonalnego i świadomego namysłu.

W przypadku natomiast świadomego umysłu, używane jest pojęcie intencji. Współcześnie mówiąc o intencji zakłada się, że ma się na myśli teorię przedmiotów świadomości zajmującą się relacjami między aktami świadomości²³ (lub samymi podmiotami tych aktów) a przedmiotami tych aktów. Jakkolwiek już w średniowieczu sformułowano zagadnienie intencjonalności, powstanie współczesnych teorii intencjonalności wiąże się z poglądami Franza Brentano. Według Brentano fenomen jest intencjonalny, jeśli jest „skierowany na coś „ lub „odnosi się do czegoś”. Rzecz jasna, przestrzenne i językowe metafory można wyeliminować na rzecz dosłownego użycia terminu „o”, co prowadzi do następującego sformułowania: fenomen jest intencjonalny jeśli jest „o” czymś.

Ponieważ kierunek naszego postępowania warunkowany jest oceną określonego stanu rzeczy w wymiarze pozytywnym lub negatywnym, rozważania o intencji pozytywnie lub negatywnie „skierowanej na coś” lub „odnoszącej się do czegoś” prowadzą nas do jeszcze głębszych rozważań na temat ocennego (poznawczego) aspektu emocji, związanej ze szczęściem, czy dobrostanem psychicznym.

3.2. Szczęście, dobrostan psychiczny, użyteczność

Dusza

Quo vadis domine? Prawie każdy wybrałby jedną z poniższych odpowiedzi: (i) podążam ku radości, szczęściu lub (ii) danej wartości. Niewątpliwie spełnienie się człowieka na wybranej drodze sprzyja jego dobrostanowi psychicznemu. Przez wieki, przede wszystkim na gruncie religii, „dobre życie” utożsamiane było najbardziej fundamentalnie rozumianymi wartościami lub wręcz wymiarem transcendentnym. Niezależnie czy mamy do czynienia z hinduizmem, buddyzmem, islamem czy chrześcijaństwem w ujęciu monistycznym ostatecznie sprowadzającym się do pojęcia „Jedności” a w ujęciu dualistycznym do innej konkretnej wartości/bytu o innej ontologii (jak dusza) niż cała reszta.

²³ Zwykle mówi się o zachowaniach intencjonalnych i nieintencjonalnych. Ten rozdział oczywiście nie jest rozłączny, np. O'Shaughnessy (1998) mówi o zachowaniu subintencjonalnym, które wprawdzie nie jest poprzedzone świadomą intencją, ale inaczej niż odruch może być zatrzymane wedle naszej woli. Przykładem może być wystukiwanie stopą rytmu piosenki albo poruszanie ustami podczas czytania. Searl (1983) wprowadził natomiast pojęcie działań nierozmyślnych, które zgodne są z intencją, ale wykonywane niejako przez przypadek.

Filozofia starożytna także wiązała dobrostan psychiczny z pojęciami fenomenalnymi i fundamentalnymi. Dla Platona, Arystotelesa, a zwłaszcza dla stoików, szczęście (dobrostan) utożsamiane z eudajmonią mało ma wspólnego z przyjemnością, czy jakimikolwiek doznaniem emocjonalnymi. Ważne staje się to co warte jest naszych starań, czyli to co jest zgodne z naszą naturą (z dajmonem czy z „prawdziwym Ja”) i przez to pozwala nam najlepiej spożytkować tkwiący w nas potencjał. Szczęściem jest życie autentyczne, podkreślające podstawowe cnoty, zapewniające samorealizację. Niekoniecznie przy tym życie przyjemne.

Obliczanie

W przeciwieństwie jednak od powszechnie dominującego do poł. XX w. poglądu o istnieniu jednostkowej duszy reprezentującej w wymiarze duchowym człowieka, dość szybko, bo jeszcze w Starożytności zaczął kształtować się odmienny od eudajmonistycznego i bardziej „przyziemny” pogląd na naturę szczęścia. Tradycja hedonistyczna (gr. ἡδονή, hedone, "przyjemność", "rozkosz") reprezentowała stanowisko, że działanie ludzkie jest motywowane wyłącznie chęcią uzyskania przyjemności i uniknięcia nieprzyjemności. Zgodnie z tym poglądem przyjemność i rozkosz uznaje się za najwyższe dobro i cel życia, główny motyw ludzkiego postępowania. Unikanie cierpienia i bólu jest głównym warunkiem osiągnięcia szczęścia. Podejście hedonistyczne ma długą tradycję. Sprowadzenie dobrostanu do zmysłowej przyjemności było zasadniczą tezą etyki Arystypa z Cyreny, ucznia Sokratesa. Ten dość prymitywny, czysto zmysłowy hedonizm został z czasem uszlachetniony. Dla utylitarystów miarą szczęścia było zaspokojenie ludzkich pragnień, nie ograniczonych do pragnień sensualnych. Podobnie dla Jeremy'ego Bethama i późniejszych utylitarystów, którzy w maksymalizowaniu indywidualnych przyjemności i zaspokajaniu innych indywidualnych potrzeb upatrywali główny warunek budowania dobrego społeczeństwa („jak najwięcej jak najszcześniejszych”). Współcześni badacze jakości życia o orientacji hedonistycznej przyjmują w ślad za utylitarystami szeroką miarę subiektywnej użyteczności, uwzględniającą zarówno bilans przyjemności i nieprzyjemności, jak i satysfakcję z realizowanych dowolnych celów w dowolnej dziedzinie.²⁴

Realizacja tych celów ściśle związana była z ich użytecznością. Dlatego też bardzo ważnym dla naukowej refleksji nad ideami hedonistyczno-utylitarystycznymi było powstanie ekonomii, dla której pojęcie użyteczności ma fundamentalne znaczenie. Ekonomia jako nauka ukształtowała się dopiero w drugiej połowie XVIII w. jako nauka zajmująca się problematyką gospodarowania (decyzji, wyborów)²⁵, a dokładniej optymalnego dysponowania ograniczonymi zasobami w warunkach na ogół niepełnej informacji. W sytuacji ograniczonych zasobów uczestnicy rynkowej gry ekonomicznej muszą dokonywać określonych wyborów. Czynnikiem determinującymi te decyzje, a przynajmniej w większym lub mniejszym stopniu na nie wpływającymi, są potrzeby. Miarą determinowania konkretnego wyboru jest jego użyteczność.

Rozwój teorii użyteczności przyczynił się do lepszego zrozumienia podejmowania decyzji, percepcji ryzyka, poziomów cen i wielu innych zagadnień ekonomicznych. Jednocześnie jest to pole, na gruncie którego powstało wiele wciąż niewyjaśnionych problemów badawczych kluczowych m.in. dla zrozumienia zjawisk zachodzących w świecie ekonomii i finansów jak chociażby problem trudnej do wytłumaczenia nadwyżki zysków z rynku akcji nad rynkiem obligacji.²⁶

W dużej mierze, do rozwoju teorii użyteczności przyczyniały się kolejne paradoksy związane z rozumieniem tego pojęcia. Jeszcze w starożytności możemy odnaleźć ślady refleksji nad pojęciem użyteczności. Arystotelesa wskazywał na paradoks wody i diamentów. Paradoks ten wskazuje na to, że

²⁴Diener E., Suh, E.M., Lucas, R.E. Smith, H.L. (1999). Subjective well being: three decades of progress. Psychological Bulletin, 125, 276-302, Czapliński J., Spotkanie dwóch tradycji eudajmonizmu i hedonizmu, Psychologia pozytywna, Warszawa 2008

²⁵ Takie rozumienie ekonomii wywodzi się ze starożytnej Grecji (V-IV w. p.n.e.) od Ksenofonta i Arystotelesa (zob. np. [Marciniak, 1999].

²⁶ Maryniak P. Ewolucja teorii użyteczności

woda pomimo, że jest niezbędna do życia jest dobrem tanim, podczas gdy diamenty, bez których z powodzeniem możemy się obejść są drogie. Próbę rozwiązania tego paradoksu podjął jeden z twórców ekonomii A. Smith, który twierdził, że paradoks wynika z tego, że wartość możemy rozumieć jako użyteczność, czyli wartość użytkową danego dobra oraz jako jego siłę nabywczą czyli wartość

wymienną. Dobra charakteryzujące się dużą wartością wymienną posiadają niską wartość użytkową i na odwrót. Toteż woda, charakteryzująca się dużą wartością użytkową ma jednocześnie niską wartość

wymienną, podczas gdy diamenty o niskiej wartości użytkowej mają wysoką wartość wymienną.²⁷ Klasycznym paradoksem, który najczęściej przywoływany jest w literaturze przedmiotu jest jednak tzw. paradoks petersburski. Zgodnie ze sformułowaną jeszcze przez Pascala ważnym kryterium wyboru w warunkach ryzyka jest tzw. zasada maksymalizacji oczekiwanej wartości. Oczekiwana wartość zdefiniowana jest jako:

$$EV = \sum p_i V(o_i)$$

gdzie: EV - wartość oczekiwana, p_i - prawdopodobieństwo wystąpienia i -tego zdarzenia, $V(o_i)$ - wartość i -tego zdarzenia.

Powyższa zasada nie może jednak być uznana za racjonalne kryterium wyboru w warunkach ryzyka. Zwrócił na to uwagę N. Bernoulli. We wspomnianym właśnie paradoksie petersburskim zadał on bowiem bardzo ważne pytanie: jak dużo powinien zapłacić potencjalny uczestnik za możliwość wzięcia udziału w grze polegającej na kolejnym rzucaniu monetą tak długo, dopóki nie wypadnie reszka (gdy wypadnie – gra zostaje zakończona). Uczestnik gry otrzymuje wypłaty według następującej zasady: jeżeli gra kończy się na rzucie $(n+1)$, to gracz otrzymuje n^2 dukatów. Jak łatwo policzyć, wartość oczekiwana tej gry wynosi nieskończoność, zatem zgodnie z zaproponowaną przez Pascala zasadą, potencjalny uczestnik powinien być skłonny zapłacić za udział nieskończenie wiele. Jednak wiadomo, że żaden gracz tak nie uczyni i nikt nie zarzuci mu, że postępuje nieracjonalnie. Krewny Nicolasa, Daniel Bernoulli wskazał, że maksymalizacja wartości oczekiwanej nie może być racjonalnym kryterium wyboru w warunkach ryzyka. Zamiast niej zaproponował on zasadę maksymalizacji oczekiwanej użyteczności. We wzorze Pascala oceniając atrakcyjność działania w warunkach niepewności, powinno się mnożyć prawdopodobieństwo nie przez wartość konsekwencji, a przez ich użyteczność. Zasada ta dopuszcza więc indywidualny stosunek do ryzyka. Pomimo tego jednak zasada ta wymaga od decydenta zachowania daleko posuniętej konsekwencji w jego preferencjach.

Formalne rozwiązanie tego problemu zostało natomiast zaproponowane przez Daniela Bernoulliego w opublikowanym w 1738 roku dziele pt. Specimen Theoriae novae de Mensura Sortis. W tym dziele Bernoulli postuluje, że osoba postawiona przed tak postawionym problemem ocenia taką wartość nie na podstawie ceny, ale użyteczności. Praca Bernoulliego pozwoliła nie tylko rozwiązać ważny problem

rachunku prawdopodobieństwa, ale położyła również fundamenty pod dzisiejsze rozumienie pojęć takich jak użyteczność czy ryzyko. Jednak w tamtych czasach bardziej niż z raczkującą ekonomią praca

Bernoulliego miała więcej wspólnego z hazardem. Dlatego dopiero w XX w. koncepcje Bernoulliego znalazły swoje miejsce w głównym nurcie ekonomii.

Refleksję nad pojęciem użyteczności rozwijali także Gossen, Jevon, Edgeworth'a, Fisher wskazując na kolejne problemy związane z jej pomiarem.

Brak możliwości efektywnego mierzenia całkowitej użyteczności przy jednoczesnym braku

²⁷ Jak słusznie jednak zauważa Stiegler (1950) argumenty Smitha miały bardziej charakter osądu niż logicznego dowodu. Smith nie uzasadnił odwrotnej proporcjonalności wartości użytkowej i wymiennej dóbr uznając to po prostu za fakt. Wyjaśnienie tego paradoksu zostało przeprowadzone dopiero w drugiej połowie XIX wieku za sprawą tzw. rewolucji marginalistycznej.

zapotrzebowania na jej mierzenie dla potrzeb analizy ekonomicznej spowodował, że oczy ekonomistów z użyteczności kardynalnej skierowały się na użyteczność porządkową. Zasada użyteczności stała się niejako podstawą ekonomii jednak w sposób odmienny niż postulowali to utylitaryści. Na mocy rewolucji ordynalistycznej użyteczność przestała być wartością normatywną. Nadal zakładano, że jednostki w swoich działaniach w sposób intuicyjny posługują się maksymalizacją użyteczności. Brak możliwości obiektywnego zmierzenia jej w sensie kardynalnym oznaczał jednak, że niemożliwe jest opieranie o nią zasad działania systemu ekonomicznego czy polityki gospodarczej. Powrót do użyteczności całkowitej nastąpił dopiero kilkadziesiąt lat później. W pracy „Theory of Games and Economic Behavior” ustanowili podwaliny dla działu matematyki zajmującym się strategicznym działaniem w warunkach konfliktu i kooperacji, czyli teorii gier, ale także przedstawili teorię oczekiwanej użyteczności. Von Neumann i Morgenstern zauważyli, że skazanie użyteczności kardynalnej „na banicję” przez ekonomistów było prawdopodobnie przedwczesne, a problem mierzalności może być w swojej naturze podobny do fizycznego mierzalności ciepła, który również początkowo oparty był o intuicyjne odczucia, a rozwiązany został wraz z wprowadzeniem temperatury, która stanowi miarę bardzo specyficzną. Von Neumann i Morgenstern wykazali, że w przypadku jeśli preferencje jednostki spełniają określone warunki to mierzenie użyteczności jest możliwe. Te warunki to aksjomat zupełności, przechodniości i ciągłości. Mierzalność użyteczności opierała się na porównaniu różnic w wartości związanej z wyborem dobra preferowanego w drugiej kolejności a wyborem w loterii pomiędzy dwoma skrajnie preferowanymi dobrami (jako pierwsze i ostatnie w kolejności). W ten sposób zaprezentowana kilkadziesiąt lat wcześniej kategoria wartości oczekiwanej została wreszcie usystematyzowana. Bardzo istotnym było także wprowadzenie na stałe do ekonomii rygoru matematycznego, co na trwałe odmieniło jej obraz.

Naturalny wydawało się więc rozszerzyć modele oczekiwanej użyteczności przez włączenie do analizy nie tylko subiektywnej oceny dotyczącej wartości konsekwencji możliwych opcji wyboru, lecz także subiektywnego traktowania prawdopodobieństw, związanych z tymi konsekwencjami. Klasycznym modelem, który rozwinął tę ideę, jest model subiektywnie oczekiwanej użyteczności (subjective expected utility, SEU), którego fundamenty znajdziemy już w oryginalnej pracy Ramseya z 1929 roku. Podobnie niezależnie zrobił to Bruno de Finetti (1937). Obaj twierdzili, że subiektywne prawdopodobieństwa mogą być wnioskowane z obserwacji ludzkich zachowań. Zaproponowali wyprowadzenie stopnia pewności, co do prawdziwości jakiegoś stwierdzenia, z analizy zachowań decydentów podczas podejmowania zakładów i wyliczenie tego stopnia jako prawdopodobieństwa. Pełny model wyboru w warunkach niepewności, który uwzględniałby subiektywne prawdopodobieństwa został przedstawiony w słynnej pracy L. Savage'a: The Foundation of Statistics (1954). Savage zsyntezował osiągnięcia Ramsey'a i Finetti'ego oraz model oczekiwanej użyteczności von Neumanna-Morgensterna. Wyprowadził nową, analityczną strukturę oraz warunki konieczne i istotne na istnienie subiektywnych prawdopodobieństw oraz na wyjątkowe ich połączenie z użytecznościami wyników dostępnych opcji wyboru. Scharakteryzował indywidualny wybór jako maksymalizujący subiektywnie oczekiwaną użyteczność.

Wydawać by się mogło, że tym samym ekonomiści uzyskali kompletny model ludzkiego zachowania. W rzeczywistości jednak teoria ta okazała się błędna w zderzeniu z wynikami empirycznymi. Już w 1953 (czyli na rok przed pierwszym wydaniem The Foundation of Statistics) na łamach czasopisma Econometrica francuski ekonomista Maurice Allais opublikował słynny paradoks, który przedstawiał sytuację decyzyjną odbiegającą od zaproponowanego przez von Neumanna-Morgensterna modelu. Paradoks ten pokazywał, że decydenci w sytuacji wyboru pomiędzy opcjami o dokładnie takich samych użyteczności oczekiwanej skłonni byli wybierać różne opcje, co wprost przeczyło przyjętej racjonalności. W latach 50. i następnym psychologowie opisali mnóstwo kolejnych odstępstw od teorii EU, charakteryzujących zachowania ludzi. Nie byli jednak w stanie zaproponować alternatywnej teorii a model ewidentnie należało uelastyczyć. Uczynili to dopiero Kahneman i Tversky (1979), formułując teorię perspektywy. Pokazali najpierw kilka głównych efektów wskazujących, jak model EU systematycznie zawodzi jako model deskryptywny. Dwa z tych efektów dotyczące nieliniowego traktowania prawdopodobieństw są pewną modyfikacją słynnego paradoksu Allais (1953).

Warto zauważyć, że problematyka perspektywy ma wiele wspólnego z problemem ramy, o którym była mowa przy okazji uwag na temat Umysłu i wiedzy. Podobnie jak trudnym okazuje się symboliczne opisanie funkcjonowania umysłu tak trudnym okazuje się również opis użyteczności oczekiwanej i kardynalnej. Podobnie jednak jak w przypadku reprezentacji umysłowych (Umysłu i wiedzy) problemy te wynikają ze złożoności umysłu i podobnie nie wykluczają możliwości coraz bardziej dokładnego ich odtwarzania czyli także modelowania wykonywania określonych wyborów.

Ucieśnienie (i uduchowanie?)

Rozwój ekonomii i handlu w oczywisty sposób związany był z narodzinami rewolucji przemysłowej. Towarzyszący jej gwałtowny rozwój techniki i nauki powodował zmiany w tym, jak ludzie postrzegali świat. Podobnie jak wizja człowieka obliczeniowego dla wielu była nie do zaakceptowania, tak wizja człowieka, którego jedynym celem była jego własna korzyść była dla nich wręcz odrażająca. Zachodzące zmiany i towarzyszący im niepokój coraz bardziej sprzyjały temu aby człowiek "spojrzał wewnątrz samego siebie". Zmiany te w końcu powstaniem nowej nauki - psychologii. Powstanie psychologii można w szerszej perspektywie można wiązać z historycznie postępującą emancypacją nauk przyrodniczych od filozofii. Metody naukowe, które mniej czy bardziej rygorystycznie stosowała psychologia coraz śmielej przekraczały granice zarezerwowane dotychczas dla filozofii. Za ojców psychologii jako odrębnej dziedziny naukowej i dyscypliny akademickiej wyodrębnionej z filozofii uznaje się Wilhelma Maximiliana Wundta (1832-1920) i Williama Jamesa (1842-1910). Zygmunta Freuda był z kolei ojcem psychologii analitycznej (psychoanalizy). Psychoanaliza jest zarówno teorią ludzkiego umysłu, jak i praktyką terapeutyczną. Zapoczątkował ją Zygmunt Freud między 1885 a 1939 rokiem i do dzisiaj jest ona rozwijana przez psychoanalityków na całym świecie. Psychoanaliza ma cztery główne obszary zastosowania: jako teoria funkcjonowania umysłu, jako metoda leczenia problemów psychicznych, jako metoda badawcza, jako sposób patrzenia na zjawiska kulturowe i społeczne. Freud wyraźnie rysuje dynamikę aparatu psychicznego z perspektywy ID przez program zaspokajania rozkoszy. Dlatego samozachowanie ego może następować tylko pod znakiem indywidualnej ekonomii libido. Psychoanaliza opiera się więc ewidentnie o paradygmat hedonistyczny. Hedonizm ten jest jednak naznaczony pesymizmem. Program rozkoszy jest bowiem nieustannie zagrożony niepowodzeniem i jeśli w ogóle udaje się go realizować, to jedynie w sposób bardzo ograniczony.

Temu dość pesymistycznemu obrazowi, także na gruncie rodzącej się psychologii, biernie wspieranej przez behawioryzm i później kognitywizm²⁸, próbowała zaradzić psychologia analityczna Carla Junga, psychologia postaci, psychologia egzystencjalna czy psychologia humanistycznej. Ujęcia te ukierunkowane były na bardziej fundamentalne i trwałe wartości niż na proste przyjemności i zmienne cele.

Psychologia pozytywna zrodziła się pod koniec XX wieku w Stanach Zjednoczonych jako, z jednej strony protest wobec psychologii negatywnej oraz jako reakcja na pogarszanie się wskaźników społecznych w najwyżej rozwiniętych gospodarczo krajach, zwłaszcza w Stanach Zjednoczonych. Ważne są także jednak naukowe ambicje psychologii pozytywnej. Jeśli z jednej strony psychologia pozytywna nawiązuje do psychologii humanistycznej, wykazując związki z fenomenologią, z drugiej psychologia pozytywna odcina się całkowicie od metafizyki i fundamentalizmu. Teleologia psychologii pozytywnej to przede wszystkim nauka o celach a nie nauka o wartościach. Cele psychologii pozytywnej nie mają nic wspólnego z fenomenalnym opisem jakiejś bardziej fundamentalnej warstwy umysłu (lub rzeczywistości). Psychologia pozytywna stara się integrować dwie tradycje podejścia do problematyki dobrostanu: hedonistyczną: niech każdy z nas będzie jedynym i ostatecznym sędzią we własnej sprawie, zdając się w ocenie i kształtowaniu swego Życia na osobiste odczucia przyjemności-przykrości, spełnienia i niespełnienia, dobra i zła, oraz eudajmonistyczną: człowiek nie zawsze wie, co w dłuższej perspektywie jest dla niego dobre, a co złe, nie zna w pełni swojej natury, potrzeb i możliwości, są przyjemności sprzyjające pełniejszemu wykorzystaniu potencjału ludzkiego.

²⁸ bardziej zajmującymi się procesami uczenia się niż odkrywaniem lub uzasadnianiem fundamentalnych wartości - co jednak w dalszej części tej pracy z tych pozycji konsekwentnie czynię

Prominentnymi reprezentantami psychologii pozytywnej byli wcześniej wspomniani David Kahnemann i Amos Tversky.

Historia współpracy Kahnemana i Tversky'ego sięga wczesnych lat siedemdziesiątych. Pierwszym programem badawczym, który przyniósł im sławę, była seria badań eksperymentalnych dotycząca wnioskowania probabilistycznego, znana jako heurystyki i zniekształcenia (Heuristics and Biases). Inspiracją dla tego programu było kilka zadziwiających efektów zaobserwowanych w laboratorium psychologicznym. Jednym z najsłynniejszych wyników takich eksperymentów, który wzbudził wiele gorących dyskusji i inspirował szereg badań z zakresu kognitywistyki, był tzw. problem Lindy. Po usłyszeniu opisu młodej kobiety o liberalnych poglądach, badani twierdzili, iż bardziej prawdopodobne jest, iż mamy do czynienia z urzędnikiem bankowym o poglądach feministycznych, niż, że mamy do czynienia z urzędnikiem bankowym. Jest to oczywiste pogwałcenie praw logiki formalnej i teorii zbiorów, gdyż zbiór feministycznie nastawionych urzędników bankowych jest podzbiorem zbioru urzędników bankowych. Innym zaskakującym rezultatem była np. opinia badanych studentów, że w rodzinach z sześciorgiem dzieci, w których jest trzech chłopców (C) i trzy dziewczynki (D), sekwencja urodzin CCCDDD jest mniej prawdopodobna niż DCCDCD, choć z rachunku prawdopodobieństwa wiemy, że wszystkie tego typu sekwencje są jednakowo prawdopodobne. Niedoskonałe zasady wnioskowania probabilistycznego zostały nazwane heurystykami, a systematyczne błędy – zniekształceniami. Przykładem heurystyki jest dostępność – szacowanie prawdopodobieństwa jakiegoś zdarzenia na podstawie tego, jak łatwo wydobywamy z pamięci jego instancje – dlatego ludzie uważają, że liczba zabójstw jest wyższa niż liczba samobójstw (przypadki zabójstw są bardziej nagłaśniane przez media i stąd lepiej zapamiętywane), choć faktycznie jest odwrotnie. W roku 1982 Kahneman i Tversky wydają razem z Paulem Slovicem kolekcję artykułów Judgment under Uncertainty: Heuristics and Biases, która podsumowywała kilkanaście lat badań nad heurystykami i zniekształceniami. Praca Heuristics and Biases wywarła duży wpływ na naszą świadomość istnienia silnych i systematycznych zniekształceń poznawczych – od 1982 roku była kilkunastokrotnie wznawiana w niezmienionej formie przez Cambridge University Press (w niektórych latach nawet dwukrotny reprint). W roku 2002 pod redakcją D. Kahnemana, T. Gilovicha i D. Griffina ukazała się praca Heuristics and Biases: The Psychology of Intuitive Judgments przedstawiająca zbiór badań inspirowanych programem heurystyk i zniekształceń z ostatnich 20 lat.

Drugim słynnym programem badawczym, który był zapewne głównym powodem przyznania Nagrody Nobla dla Kahnemana, była teoria perspektywy (prospect theory). W 1979 roku w Econometrica ukazuje się artykuł obu uczonych przedstawiający nową teorię zachowania ludzi w sytuacjach ryzyka. Teoria perspektywy wyjaśniała w prosty sposób wiele anomalii ekonomicznych, czyli zachowań, niezgodnych z ekonomicznie rozumianą racjonalnością (ekonomiści uznają zachowanie za racjonalne, jeśli wybory jednostki są w zgodzie z tzw. aksjomatami teorii oczekiwanej użyteczności). Model Kahnemana i Tversky'ego zakładał, że decydenci oceniają potencjalne wyniki ryzykownych decyzji finansowych biorąc pod uwagę nie ich wpływ na całkowity majątek decydenta, ale na aktualny punkt odniesienia, którym jest zwykle zerowy zysk/strata. Ludzie, zgodnie z zasadą „lepszy wróbel w garści niż gołąb na dachu”, niechętnie ryzykują w sferze zysków, ale często ryzykują w sferze strat – aby uniknąć mniejszej straty ryzykują poniesienie większej np. wdając się policjantem w dyskusję po to, by uniknąć mandatu. Zauważono również, że smutek z powodu utraty sumy X wydaje się być ponaddwukrotnie silniejszy niż radość z otrzymania sumy X, co zostało nazwane awersją do strat (loss aversion). Wydarzenie bardzo mało prawdopodobne otrzymują nieproporcjonalnie dużą wagę w procesie decyzyjnym, podobnie jak wydarzenia, które zdarzą się na pewno w stosunku do tych, które zdarzą się prawie na pewno. Bardzo ważną implikacją teorii perspektywy jest to, że silny wpływ na decyzję ma sformułowanie problemu decyzyjnego, np. to, czy powiemy „ten plan ratunkowy pozwala uratować 100 osób z pięciuset” zamiast „przy tym planie ratunkowym 400 z 500 osób nie przeżyje”. Model Kahnemana i Tverskyego pozwala przewidywać i wyjaśniać zachowania zwykłych konsumentów, jak i ekspertów (np. lekarza stawiającego diagnozę), inwestorów giełdowych czy za rządów przedsiębiorstw. W 1999 roku, czyli już po śmierci Tversky'ego, ukazuje się zbiór artykułów: Choices, Values and Frames, który zawiera ponad trzydzieści prac inspirowanych teorią perspektywy.

Ostatni rozbudowany projekt badawczy, którym Kahneman zajął się już po śmierci Amosa Tversky'ego

dotyczy procesów oceniania doświadczeń przyjemnych i nieprzyjemnych i jest znany jako hedonic psychology. To, w jaki sposób ludzie zapamiętują przyjemne i nieprzyjemne przeżycia, okazuje się kolejnym wyzwaniem dla potocznych wyobrażeń i ekonomicznych postulatów racjonalności. Okazuje

się choćby, że z dwóch bolesnych zabiegów medycznych, ten, który trwa dłużej, może być w pewnych

warunkach zapamiętany jako mniej nieprzyjemny niż podobny zabieg trwający krócej.

Dzięki Danielowi Kahnemanowi i Amosowi Tversky'emu posiadamy usystematyzowaną wiedzę na temat niedoskonałości naszych wnioskowań statystycznych, zarówno tych codziennych, jak i dokonywanych w formie poważnych ekspertyz. Potrafimy trafniej wyjaśnić dziwaczne zachowania zwykłych konsumentów i wielkich przedsiębiorstw. Lepiej rozumiemy, dlaczego ludzie kupują losy na loterię, a także, czemu premia za ryzyko na rynkach finansowych jest tak wysoka. Zapoczątkowany przez nich nurt badań z pewnością zaowocuje jeszcze szeregiem intrygujących odkryć o niebagatelnym znaczeniu praktycznym.

Dominującym kierunkiem w psychologii szczęścia jest zapewne kierunek hedonistyczno-utylitarystyczny, trudno bowiem nie odczuwać, w większości przypadków, oczywistego związku między szczęściem a przyjemnością, czy spełnianiem celów. Kierunek ten ma wymiar praktyczny, szczególnie dla takich nauk jak ekonomia, bowiem przyjemności i cele dają się jakoś mierzyć (według zadanego kryterium użyteczności), o co o wiele trudniej w przypadku wiązania wiązania szczęścia z określonymi wartościami, sformułowanymi w najogólniejszy sposób.

Jednak trudno także nie zauważyć, że wartości te wcale nie muszą stać w opozycji do prostych przyjemności i celów, a mogą z nich wprost wynikać. Takie wskazówki daje nam ewolucyjna teoria gier, która powstała na podwalinie teorii gier stworzonej przez wspomnianych wcześniej Morgensterna i von Neumanna albo istnienie np. neuronów lustrzanych²⁹. Trudno także nie zauważyć, że poziom szczęścia ostatecznie zależy po prostu od uwarunkowania skojarzeń uruchamiających określone procesy biochemiczne prowadzące do odczucia pozytywnych doznań. Odpowiednie treningi sieci neuronowych w duchu psychoterapii behawioralno-kognitywnej jest także sposobem na poprawę samopoczucia. Podobnie jest z medytacją lub kontemplowaniem określonych wartości, często podejmowanych przez religie i filozofie. Taką inspiracją do zmiany określonych skojarzeń może być wizja świata i nas samych dość jednoznacznie wynikająca ze znanej już nam wiedzy: **“n.p. jesteśmy bardziej częścią Świata niż jakimś autonomicznym bytem, Świat jest jedną całością a to co wspólne i nas spaja jest o wiele ważniejsze od nas samych.”**

;

3.3. Wartość, dobro, normatywność i użyteczność.

Dusza

Obecny w religiach i filozofii świat wartości i dobra, jeszcze mocniej niż szczęścia, odcinał się od cielesnej przyjemności próbując ugruntować inną niż wspólną im i materii ontologię.

Jak omówione zostało to już wcześniej w wcześniejszym rozdziale, już starożytni Grecy zauważyli, że człowiek składa się z czegoś więcej niż tylko z ciała. Heraklit zauważył, że człowiek posiada w sobie duszę, która jest przyczyną ciągłych zmian.

Po nim przyszedł Platon, twierdząc, że zasadniczo człowiek jest cały duszą, która została zesłana do ciała w ramach kary za przewinienia, które mogła popełnić przebywając w świecie idei (czyli świecie realnym, prawdziwym; Platon wszystko, co się wiązało z ciałem, światem materialnym, uważał za nędzny cień świata idei). Nazwał on ciało grobem duszy, z którego może ona się wydostać jedynie poprzez cnotliwe życie. Pierwotnie poglądy Platona były w kwestii rozdziału dobra i psychologicznego szczęścia dość radykalne, rozdział pomiędzy dobrem doczesnym a dobrem idealnym był dla niego

²⁹ powrócę do tych przykładów w części dotyczącej Wartości, dobra i piękna

zasadniczy. Później uznał on jednak, że dobra realne za niezbędny szczebel do osiągnięcia dóbr idealnych. Etyka Platona ostatecznie zakładała, że dobra doczesne są szczeblem do dóbr idealnych. Teorię Platona poprawił jego uczeń Arystoteles, twierdząc, że dusza i ciało są ze sobą nierozdzielalne, tworząc razem byt jakim jest człowiek. I to właśnie dusza jest tym elementem, który nadaje ciało kształt. Jest jego formą materialną, i jako taka powołuje ciało do życia, jest jego pierwszym aktem konstytuującym, organizatorką życia ciała człowieka. Arystoteles z kolei szczęścia i wartości (dobra) doszukiwał się w eudajmonii. Tą eudajmonię Arystoteles nie cechował jednak antynaturalizm. Przesuwając szczęście bardziej ku wartościom i oddalając je jednocześnie od hedonistycznej przyjemności wciąż widział on zarówno materię i duszę jako właściwości jednego świata. Podobnie jak wpisane w ich naturę cele. Takie jaką właśnie dla człowieka była eudajmonia, czyli rozumne życie. Dla starożytnych filozofów zasadniczo szczęśliwe życie było równoznaczne z wartościowym życiem, jednak już wtedy narodził się kierunek, który upatrywał szczęścia i wartościowego życia w znacznie prostszych przyjemnościach.

Obliczenie ??

Nieco zabawnym jest fakt, że w odmiennie niż w przypadku Umysłu, Świadomości czy Szczęścia, zanim w XXw nauka na dobre zaczęła kwestionować nasze wyobrażenie o świecie i przede wszystkim nas samych, w przypadku Wartości i Dobra nauka, w szczególności logika zdań (jako dziedzina matematyki), najpierw stanęła po "tradycyjnej" stronie. Współtwórcy nurtu znanego filozofia analityczna, George Edward Moore wraz z Bertrendem Russellem sprzeciwiali się idealizmowi, reprezentowanemu w Wielkiej Brytanii przez neoheglistów Francis Herberta Bradleya i Johna Ellisa McTaggarta. Filozofia analityczna uprawiana była później przez Ludwiga Wittgensteina, Koło Wiedeńskie (R. Carnap), filozofów języka potocznego (Austina, Ryle'a, II Wittgensteina) oraz filozofów amerykańskich (S. Kripkego, H. Putnama, D. Davidsona, J. Searle'a). Źródłem filozofii analitycznej był tzw. zwrot lingwistyczny (linguistic turn), który nastąpił pod wpływem przełomu, jaki dokonywał się w logice (wydanie Principia mathematica B. Russella i A. N. Whiteheada). Filozofią analityczną cechował więc zatem: traktowanie języka jako narzędzia i przedmiotu filozofii, dbałość o precyzję argumentacji, niechęć do pospiesznych syntez i rezygnacja z budowania systemu filozoficznego na rzecz drobiazgowej analizy poszczególnych zagadnień, intersubiektywna sensowność i kontrolowalność wyników, czysto poznawcze cele filozofii. Kluczowymi metodami filozofii analitycznej była: klasyczna analiza pojęciowa, analiza logiczna, analiza lingwistyczna i analiza systematyczna.

W 1903 roku w Principia Ethica George Edward Moore wysunął brzemiennej w konsekwencje tezę: pojęcia moralne, takie jak np. „dobry”, nie mogą być definiowane za pomocą pojęć naturalnych. Zarzucił on niektórym teoriom etycznym, iż obarczone są błędem, nazwanym przez niego „naturalistycznym”. Jako jeden z twórców filozofii analitycznej był przekonany, że problemy etyczne mogą być rozwiązane na drodze analizy języka. Jego zdaniem rozumieć znaczenie jakiegoś predykatu to tyle, co znać własność, do której się on odnosi. Aby zdefiniować dany termin, należy podać rozwinięte, synonimiczne wyrażenie, które można prawdziwie orzec o własności nazywanej przez ów termin. Okazuje się jednak, iż istnieją pojęcia, których zdefiniować nie można. Takim pojęciem jest „dobry”. Moore jako dowód tej tezy zaprezentował tzw. test otwartego pytania. Test ten można zrekonstruować w następujący sposób 5 : Załóżmy, że: 1) Predykat „dobry” jest analitycznym odpowiednikiem naturalnego predykatu „N”. Zatem, na mocy definicji analityczności wiemy, że: 2) Znaczenie tezy „x jest N” jest identyczne ze znaczeniem tezy „x jest dobry”. Gdyby tak jednak było, to: 3) Ktoś, kto poważnie zadaje pytanie „Czy x, który jest N, jest również dobry?”, popełniłby błąd pojęciowy. Jednakże nie wydaje się, aby (3) było prawdą, gdyż, 4) Pytanie „Czy x, który jest N, jest również dobry?” jest pytaniem otwartym – znaczącym i niejawiącym się jako pomieszanie pojęć (np. można sensownie zapytać, „Czy czyn x, który jest pożądanym przez większą część społeczeństwa, jest również dobry?"). Zatem: 5) Nie jest tak, że predykat „dobry” jest analitycznym odpowiednikiem predykatu „N”. Zatem: 6) Na mocy konieczności pojęciowej własność bycia dobrym nie może być identyczna z własnością bycia N. W konsekwencji Moore stwierdził, że własność dobra jest samoistna, prosta i niedefiniowalna, gdyż nie można zredukować jej do żadnego kompleksu

własności złożonych. Jest ona również nienaturalna, gdyż istnieje poza czasem i przestrzenią. Ponieważ dziedziną badania naukowego mogą być tylko cechy naturalne, naturalności, zastępując przedmiot etyki przedmiotem jednej z nauk (najczęściej psychologii lub socjologii), postępują błędnie. Krytyka ta otworzyła nową kartę w dziejach metaetyki.

Ucieleśnienie

Jednak rozwój technologii i nauki w XXw., w szczególności porażki i ograniczenia klasycznych systemów regulowych, sprawiają, że coraz częściej zdajemy sobie sprawę, że takie ujęcie jest po prostu zbyt proste. Moore *de facto* arbitralnie stwierdził, że własność dobra jest samoistna, prosta i niedefiniowalna, gdyż nie można zredukować jej do żadnego kompleksu własności złożonych. Tymczasem może ona nie być ani samoistna i prosta i może być wynikiem określonych procesów uczenia się na bardziej rudymenarnych fenomenach, może być z nich złożona i tak właśnie definiowana. Tym bardziej może być ona naturalna i jak najbardziej istnieć w czasie i przestrzeni (o ile w ich ramach będziemy ujmować rzeczywistość). Moore, z powodu swojego analitycznego nastawienia, arbitralnie dokonał poważnych uproszczeń i założeń na temat charakteru własności dobra. Nawet potoczne rozumienie dopuszcza stwierdzenie, że dobre jest, że jest dobre, bo przecież czasami jest źle, że jest dobrze (jak wiemy, że dobro poprzedza zła a akurat mówimy w takim kontekście, że szczególnie istotne jest to jak będzie a nie jest aktualnie teraz). Wartość, dobro, podobnie jak szczęście czy Świadomość lub Umysł można sprowadzać do funkcji konkretnego mózgu. Można oczywiście twierdzić, że funkcje te w jakimś sensie, jak platońskie cienie reprezentują jakieś realne Wartości, ale cokolwiek chcemy o nich powiedzieć, to tylko widząc je przez pryzmat tych funkcji. Nie da się uciec od kognitywnej soczewki.

Bardzo ciekawe jest to, w jaki sposób teoria gier tłumaczy altruizm, czyli takie zachowanie organizmu, które obniża jego własne dostosowanie na rzecz wzrostu dostosowania innego gracza. Robert Aumann pokazał w 1959 roku, że w nieskończonym ciągu takich rozgrywek, współpraca może być stanem równowagi. Z kolei w 1984 roku Robert Axelrod zaprosił akademików z całego świata do uczestnictwa w turnieju dla programów komputerowych, grających w iterowany dylemat więźnia. Przesyłane programy różniły się pod względem złożoności, startowego zachowania, reakcji na działanie przeciwnika itp. Wyniki pokazały, że przy wielokrotnych rozgrywkach, egoistyczne strategie dawały średnio bardzo małe wygrane w porównaniu z bardziej altruistycznymi. Axelrod pokazał w ten sposób możliwość ewolucyjnego wykształcenia się zachowań altruistycznych z nastawionych na własny zysk, wyłącznie za pomocą selekcji naturalnej.

Modnym źródłem inspiracji dla badaczy zajmujących się czy to szczęściem czy wartościami jest odkrycie tzw. neuronów lustrzanych. Istnienie tych neuronów w oczywisty sposób kojarzy się z współodczuwaniem i empatią, te z kolei z przeciwstawianym egoizmowi altruizmem. Dzięki nim możliwa jest interpretacja obserwowanego zachowania, która odbywa się w mózgu odbiorcy poprzez swoistą symulację. Nasz mózg pozwala nam odtworzyć automatycznie zachowanie innej osoby, dzięki temu możemy odtworzyć (zmapować) jego zachowanie i uczucia. Neurony lustrzane umożliwiają mentalne "czytanie" i "odtworzenie" zachowań innego mózgu. U człowieka odpowiadają prawdopodobnie również więc za zdolność do rozpoznawania cudzych emocji i intencji wyrażanych niewerbalnie. To pierwotnie "narzędzie" (służące przewidzeniu zachowania drugiego osobnika celem osiągnięcia własnych korzyści) ostatecznie stało się źródłem empatycznych zachowań odnoszących się do wielu Wartości.

Współczesna kognitywistyka jest konsekwentnie monistyczna i naturalistyczna. Jedne fenomeny wyłaniają się z drugich w drodze zasadniczo deterministycznych procesów, a przynajmniej deterministycznych w ten sposób, że zrozumienie deterministyczne jest wystarczające do tego aby wyjaśnić procesy zachodzące w naszym Umyśle. Być może dla wielu Świat, w którym pojedynczy człowiek jest w zasadzie pozbawiony woli i własnej autonomii. Jednak poczucie, że

przede wszystkim jesteśmy części związanej i splątanej ze sobą Całości wydaje się być jeszcze bardziej fascynujące.