

KOGNITYWISTYKA

SŁOWNIKI SPOŁECZNE

seria pod redakcją

Wita Pasierbka i Bogdana Szlachty



KOGNITYWISTYKA

redakcja Józef Bremer



dotychczas ukazały się:

PEDAGOGIKA RELIGII, redakcja Zbigniew Marek i Anna Walulik

POLITYKI PUBLICZNE, redakcja Artur Wolek

ETYKA POLITYCZNA, redakcja Piotr Świercz

GEPOLITYKA, redakcja Jacek Kloczkowski

PRZEDSIĘBIORCZOŚĆ, redakcja Wit Pasierbek i Krzysztof Wach

WIELOKULTUROWOŚĆ, redakcja Bogdan Szlachta

GLOBALIZACJA I WSPÓŁZALEŻNOŚĆ, redakcja Alicja Malewska i Mateusz Filary-Szczepanik

STUDIA KULTUROWE, redakcja Leszek Korporowicz, Agnieszka Knap-Stefaniuk i Łukasz Burkiewicz

BEZPIECZEŃSTWO PUBLICZNE, redakcja Tomasz Grabowski

SPORY MORALNE, redakcja Piotr Duchliński

HUMANISTYKA WSPÓŁCZESNA, redakcja Bogusława Bodzioch-Bryła

POLITYKA HISTORYCZNA, redakcja Joanna Lubecka i Maciej Zakrzewski

SŁOWNIKI
SPOŁECZNE

KOGNITYWISTYKA

REDAKCJA

Józef Bremer

WYDAWNICTWO NAUKOWE
UNIwersytetu Ignatianum w Krakowie

KRAKÓW 2024

© Uniwersytet Ignatianum w Krakowie, 2024

Publikacja finansowana w ramach zadania zleconego
ze środków Ministra Nauki i Szkolnictwa Wyższego
UMOWA Nr MEiN/2021/DPI/178



Ministerstwo Nauki
i Szkolnictwa Wyższego

Recenzenci

dr hab. Marcin Trybulec

prof. dr hab. Urszula Żegleń

Redakcja i korekta

Klaudia Bień

Projekt okładki

Studio Photo Design – Lesław Sławiński

Typografia i łamanie

Jacek Zaryczny

ISBN 978-83-7614-622-5 (Print)

ISBN 978-83-7614-630-0 (Online)

Wydawnictwo Naukowe Uniwersytetu Ignatianum w Krakowie

ul. Kopernika 26 • 31-501 Kraków

tel. 12 39 99 620

wydawnictwo@ignatianum.edu.pl

<http://wydawnictwo.ignatianum.edu.pl>

Druk i oprawa: HORN PRINT S.J. • Poznań

Spis treści

Słowo wstępu	7
Wprowadzenie	9
ANDRZEJ DĄBROWSKI Intencjonalność	13
PAWEŁ FORTUNA, ZBIGNIEW WRÓBLEWSKI Dusza	37
EWA ODOJ Problem oddziaływania psychofizycznego	57
MAREK POKROPSKI Naturalizacja fenomenologii	77
JACEK BIELAS Ucieleśnione poznanie	97
JÓZEF BREMER Kognitywistyka społeczna	117
EWA JAGLARZ Zaburzenia mowy i języka	137
SŁAWOMIR SZTAJER Religioznawstwo kognitywne	157
MICHAŁ WIERZCHOŃ Substytucja sensoryczna	175
PIOTR KONDERAK Semiotyka kognitywna: zwrot ku znaczeniu	191

ŁUKASZ WIRASZKA	
Językoznawstwo kognitywne	207
MATEUSZ HOHOL, NINA BAŻELA	
Poznanie matematyczne	229
MAREK LECHNIAK	
Logika współczesna i jej rola w naukach ścisłych oraz humanistycznych	249
ANITA PACHOLIK-ŻUROMSKA	
Logiki niemonotoniczne w modelowaniu umysłu	271
ANDRZEJ BIELECKI	
Cybernetyka	289
ANDRZEJ BIELECKI	
Sztuczne sieci neuronowe	307
MARIUSZ FLASIŃSKI	
Sztuczna inteligencja	323
DARIUSZ ADAMEK	
Reakcje emocjonalne	341
MATEUSZ JARMUŻEWSKI	
Neurobiologiczne podstawy odpowiedzialności moralnej	363
ALEKSANDRA KUZIOR	
Technologie kognitywne	385
Wykaz Autorów	403

Słowo wstępu

W roku 2019 wśród członków zespołu z „Ignacjańskiego Forum Społecznego” zrodziła się myśl kontynuowania niegdysiejszego dzieła *Słownik społeczny* z roku 2004, w którym ponad stu uczonych zarówno z ośrodków krajowych, jak i zagranicznych, zaprezentowało w obszernych esejach dorobek nauk humanistycznych i społecznych zgromadzony do początków XXI wieku.

Nowy projekt zakrojony jest szerzej niż tamten: w liczącej ponad dwadzieścia tomów publikacji zaprezentowany zostanie stan wiedzy humanistycznej i społecznej trzeciej dekady XXI stulecia; wiedzy dotyczącej człowieka rozwijającego się w zróżnicowanych cywilizacjach, kulturach i społeczeństwach, wyznającego rozmaite religie i honorującego różne wzorce postępowania. Poszczególne tomy, podobnie jak pierwsze cztery wydane już w dwóch wersjach językowych (po polsku i angielsku), dostępne również w wersji elektronicznej, traktują o polach badawczych, które obejmują zagadnienia uznawane za szczególnie ważne w naukach humanistycznych i społecznych, związane tak z pojmowaniem człowieka i jego otoczenia społecznego, jak i spraw politycznych i publicznych czy stosunków międzynarodowych. Ich analiza uwzględniająca różne podejścia badawcze pozwala pełniej zaprezentować problemy znajdujące zasadniczo w ramach jednej z dyscyplin i znakomicie poszerzyć horyzont poszukiwań podejmowanych przez wykonawców projektu. Każdy z nich szuka „klucza interpretacyjnego”, przy użyciu którego prezentuje najistotniejsze zagadnienia kojarzone z głównymi kategoriami – kontrowersyjnymi niekiedy lub budzącymi

dyskusje wśród naukowców – wykorzystywanymi w poszczególnych polach badawczych decydujących o tytułach nadawanych poszczególnym tomom nowego a wielotomowego *Słownika społecznego*. „Klucz interpretacyjny” nie byłby istotny, gdyby hasła ogłaszane w poszczególnych tomach traktowane były „zdawkowo”, na podobieństwo haseł encyklopedycznych; staje się on istotny, jeśli opracowania mają formę rozbudowanych esejów liczących około dwudziestu stroniczynormalizowanego maszynopisu; esejów zbudowanych wedle jednolitego schematu, prezentujące rozważania o istocie analizowanej kategorii, o jej dziejach, o głównych treściach z nią związanych, i – wreszcie – o zagadnieniach, które wiążą się z nią w praktyce. Całość opracowana jest na bazie refleksji teoretycznej, wzmacnianej namysłem nad praktyczną stroną, prezentowaną w dwudziestu hasłach przygotowywanych przez polskich uczonych reprezentujących nie tylko różne ośrodki i różne dyscypliny naukowe, ale także różne „wrażliwości badawcze”. Do „klucza interpretacyjnego” włączyliśmy także katolicką naukę społeczną, bowiem nie sposób pominąć dwudziestowiekowej spuścizny i bogactwa chrześcijaństwa.

Mamy nadzieję, że lektura prezentowanego tomu zadowoli Czytelników, dając Im nie tylko okazję zapoznania się z naukowymi ujęciami ważnych problemów, z którymi zмага się współczesny człowiek i zmagają się współczesne państwa oraz społeczeństwa, ale także wgląd w trudne niekiedy zagadnienia współczesności dokonywany z perspektywy katolickiej. Mamy też nadzieję, że docenią Oni trud polskich uczonych podejmujących oryginalny namysł nad nimi, nie ograniczających się do prezentowania cudzych jedynie przemyśleń, bo świadomych wagi dorobku intelektualnego rodzimej nauki.

Redaktorzy serii
Wit Pasierbek i Bogdan Szlachta

Wprowadzenie

Od czasów Platona i Arystotelesa filozofia, a później także inne nauki poszukiwały odpowiedzi na pytanie o to, czym są stany psychiczne: myślenie, wola, emocje – i jak są one powiązane z naszym materialnym ciałem. Filozoficzne rozważania na ten temat zostały wsparte na przełomie XIX i XX wieku badaniami eksperymentalnymi w psychologii, prowadzonymi przez Wilhelma Wundta. Kolejna grupa psychologów eksperymentalnych, pozostająca pod wpływem Iwana Pawłowa i innych fizjologów, zaproponowała rewolucyjne przededefiniowanie stanów psychicznych jako obserwowalnych zachowań. Rozwijający się behawioryzm przyniósł ze sobą istotną zmianę spojrzenia na ludzką psychikę w stosunku do poprzednich perspektyw bazujących głównie na introspekcji. Negował on wartość pierwszoosobowych badań świadomego umysłu i przyjmował, że psychologia – jako dyscyplina naukowa skupiająca się wyłącznie na zachowaniach – powinna być uprawiana z trzecioosobowego punktu widzenia.

Kognitywistyka (*cognitive science*) jest dzieckiem lat 50. XX wieku, produktem czasów, kiedy psychologia, antropologia i lingwistyka na nowo redefiniowały metody i zakresy swoich badań, a informatyka i neuronauka powstawały jako odrębne dyscypliny naukowe i techniczne. Rewolucja poznawcza w psychologii była swoistą kontrewolucją skierowaną przeciwko behawioryzmowi. Do łask wróciło pytanie o status badań pierwszoosobowych. Istotne zmiany zachodziły także w innych dyscyplinach. Noam Chomsky, współtwórca gramatyki transformacyjno-generatywnej, wniósł duży wkład w rozwój psycholingwistyki i informatyki. Popularność zdobywała cybernetyka Norberta Wienera, Marvin Minsky i John McCarthy pracowali nad sztuczną inteligencją,

a Alan Newell i Herbert Simon używali komputerów do symulacji ludzkich procesów poznawczych. Alan Turing zaproponował test mający na celu określanie zdolności maszyny do posługiwania się językiem naturalnym, co miało pośrednio dowodzić opanowania przez nią umiejętności myślenia w sposób podobny do ludzkiego.

Dotychczasowe monodyscyplinarne badania naukowe nad świadomym umysłem i poznawaniem rzeczywistości zrodziły konieczność prowadzenia badań interdyscyplinarnych. Zaczęto budować maszyny, które dzięki prostemu zastosowaniu algorytmów były znacznie wydajniejsze w rozwiązywaniu skomplikowanych zadań obliczeniowych aniżeli ludzie. Rozwój technik neuroobrazowania pozwalał na obserwacje pracy żywego mózgu i skorelowanych z nim stanów psychicznych osoby. W lingwistyce i w filozofii zaczęto doceniać znaczenie i funkcjonowanie języka potocznego, tego, w którym pierwszoosobowo mówimy o sobie, o swoich stanach wewnętrznych i o stanach osób trzecich. W horyzoncie tych i innych badań naukowych pojawiła się kognitywistyka. Na pytanie o to, czym zajmuje się kognitywistyka, sztuczna inteligencja ChatGPT odpowiada:

Kognitywistyka jest interdyscyplinarną dziedziną nauki, która zajmuje się badaniem procesów poznawczych, czyli procesów związanych z przetwarzaniem informacji, myśleniem, uczeniem się, zapamiętywaniem, rozumieniem języka, podejmowaniem decyzji oraz funkcjonowaniem mózgu. Kognitywistyka łączy w sobie elementy psychologii, filozofii, nauk o komputerach, lingwistyki, antropologii, neurobiologii i innych dziedzin, aby lepiej zrozumieć naturę ludzkiego umysłu oraz inteligencji.

Precyzując pytania dotyczące poszczególnych nauk i ich roli w kognitywistyce, otrzymamy coraz dokładniejsze odpowiedzi.

Początkowo badania kognitywistyczne korzystały głównie z osiągnięć informatyki, w następnym okresie zaznaczyła się dominacja neurokognitywistyki, a w ostatnich latach dowartościowano rolę czynników społeczno-kulturowych w badaniach systemów poznawczych (powstała m.in. kognitywistyka społeczna). Zakresowo obszar badań kognitywistycznych obejmuje tematykę haseł zamieszczonych w niniejszym tomie Słowników. Chodzi o pole wspólne filozofii umysłu (*philosophy of mind*), psychologii poznawczej (*cognitive psychology*), nauk komputerowych (*computer science*), sztucznej inteligencji (*artificial intelligence*), nauk o układzie nerwowym (*neuroscience*) i innych. Interdyscyplinarny charakter kognitywistyki ukazuje się w jej metodycznym, interdyscyplinarnym podejściu do przedmiotu badań, co podkreślają autorzy poszczególnych haseł. Łączy

ono analityczne metody badań obecne w naukach humanistyczno-społecznych i naukach formalnych (np. w logice, językoznawstwie teoretycznym) z postępowaniem przyrodniczo-eksperymentalnym znanym z biologii, psychologii i neuronauk. Wszystko to w połączeniu z badaniami nad rozwojem i coraz szerszym zastosowaniem sztucznej inteligencji w naszym życiu indywidualnym, społecznym, a także w przemyśle.

Poprzedniczką badań prowadzonych w obrębie kognitywistyki była cybernetyka. Wychodziła ona z założenia, że niezależnie od tego, czy system jest maszyną, organizmem żywym czy strukturą społeczną, to procesy związane ze sterowaniem i przetwarzaniem w systemie informacji i energii mają wspólne cechy. Współczesne badania kognitywistyczne nad intencjonalnością i emocjonalnością stanów psychicznych łączą przykładowo analizy dotyczące obserwowalnych zachowań osoby w otoczeniu społecznym z poszukiwaniem zmian zachodzących w odpowiednich obszarach jej mózgu. Często obszary te trudno jednoznacznie zidentyfikować, zakłada się więc, że zazwyczaj takie i inne części anatomiczne wchodzi w ich skład. Badania nad poznaniem matematycznym wykazały, że opiera się ono na ewolucyjnie starszych, uniwersalnych kulturowo systemach wiedzy wrodzonej, które wykształcają się wcześniej w toku rozwoju osobniczego. Religioznawstwo kognitywne bada procesy i mechanizmy poznawcze zaangażowane w tworzenie, przyswajanie i przekazywanie idei religijnych, sięgając do wiedzy o ewolucji ludzkiego umysłu, wiedzy o neuronalnych podstawach doświadczenia religijnego, a także badań nad interakcją umysłu z otaczającym światem. Badania kognitywne budują rozumienie wyrażen języka naturalnego (potocznego) na lingwistycznych teoriach gramatyki, sprawdzają psycholingwistyczne hipotezy o procesach przetwarzania danych za pomocą badań eksperymentalnych. Odwołując się do neuronaukowych badań eksperymentalnych i komputerowo-lingwistycznych technik programowania (opierających się na sztucznej inteligencji), tworzą modele symulacyjne (tzw. modelowanie kognitywne).

Zakres tematyczny tomu *Kognitywistyka* pokrywa się z powyższymi, przykładowymi charakterystykami badań, stanowiąc zarazem interdyscyplinarną płaszczyznę wymiany myśli. Zawiera artykuły skupiające się na wymienionych powyżej przez ChatGPT obszarach kognitywistyki. Znajdziemy w nim hasła skupione na aspekcie filozoficznym (*Intencjonalność*, *Dusza*, *Problem oddziaływania psychofizycznego*, *Naturalizacja fenomenologii*), na naukach społecznych (*Ucieleśnione poznanie*, *Kognitywistyka*

społeczna, Zaburzenia mowy i języka, Religioznawstwo kognitywne, Substytucja sensoryczna), na lingwistyce (*Semiotyka kognitywna: zwrot ku znaczeniu, Językoznawstwo kognitywne*), na logice i matematyce (*Poznanie matematyczne, Rola i znaczenie logiki w naukach, Logiki niemonotoniczne w modelowaniu umysłu*), cybernetyce i sztucznej inteligencji (*Cybernetyka, Sztuczne sieci neuronowe, Sztuczna inteligencja*), na neuronaukach (*Reakcje emocjonalne, Neurobiologiczne podstawy odpowiedzialności moralnej*) i na nowoczesnych technologiach (*Technologie kognitywne*).

Do współuczestnictwa w jego tworzeniu zostali zaproszeni wybitni naukowcy, specjaliści z kilku krajowych ośrodków naukowo-akademickich: Uniwersytetu Jagiellońskiego, Collegium Medicum Uniwersytetu Jagiellońskiego, Akademii Górniczo-Hutniczej w Krakowie, Katolickiego Uniwersytetu Lubelskiego Jana Pawła II, Uniwersytetu im. Adama Mickiewicza w Poznaniu, Uniwersytetu Marii Curie-Skłodowskiej w Lublinie, Uniwersytetu Mikołaja Kopernika w Toruniu, Uniwersytetu Komisji Edukacji Narodowej w Krakowie, Politechniki Śląskiej w Gliwicach, Uniwersytetu Ignatianum w Krakowie oraz z Uniwersytetu w Windesheim.

Żyjemy w społeczeństwie, w którym cyfryzacja danych ma ogromny wpływ na każdy aspekt naszego życia: na to, w jaki sposób porozumiewamy się i utrzymujemy kontakty towarzyskie, jak pracujemy, uczymy się, dbamy o zdrowie i uczestniczymy w życiu politycznym, społecznym czy gospodarczym. Z jednej strony cyfryzacja obiecuje ogromne korzyści: np. lepszą opiekę zdrowotną, bardziej wydajną mobilność, efektywne używanie energii, poprawę prosperowania firm i przedsiębiorstw itp. Z drugiej strony niesie również złożone pytania, np. o możliwość nieuprawnionego dostępu do wrażliwych danych o naszym życiu i o kontrolę nad nimi. Co znaczy być racjonalną, wolną i świadomą osobą w świecie, który jako osoby dzielimy z roślinami, zwierzętami i z artefaktami będącymi wytworami sztucznej inteligencji? Na ile istotna jest wiedza pierwszoosobowa, do której często się odwołujemy przy zagadnieniach etycznych? Pojawiają się również pytania o możliwości identyfikowania wiedzy rzetelnej i o poszukiwanie prawdy wśród zalewu informacji. Aby odnieść się do tych i podobnych kwestii, konieczna jest wspomniana współpraca naukowców w obszarze kognitywistyki oraz jej dalszy i systematyczny rozwój.

Redaktor tomu
Józef Bremer

Andrzej Dąbrowski

Uniwersytet Komisji Edukacji Narodowej w Krakowie

<https://orcid.org/0000-0002-3998-9608>

<https://doi.org/10.35765/slowniki.447>

Intencjonalność

Streszczenie

DEFINICJA POJĘCIA: Intencjonalność to ukierunkowanie na coś. Jest to cecha naszych aktów mentalnych, świadomości i umysłów w ogóle, która polega na ukierunkowaniu ich na pewien obiekt: myślenie jest myśleniem o czymś, świadomość – świadomością czegoś. Od tak rozumianej intencjonalności należy odróżnić intencjonalność kolektywną (*we-intentionality*), która odpowiada za działanie grupy jako pewnej całości wedle wspólnych intencji.

ANALIZA HISTORYCZNA POJĘCIA: Problematyka intencjonalności była już obecna w filozofii starożytnej, m.in. w myśli Parmenidesa, Platona czy Arystotelesa, filozofii scholastycznej i nowożytnej. Do filozofii współczesnej intencjonalność jako cechę odróżniającą zjawiska psychiczne od fizycznych wprowadził Franz Brentano. Późniejsze badania nad tą kategorią, podjęte przez przedstawicieli różnych szkół i nurtów, doprowadziły do różnych interpretacji i ścieżek myślenia.

UJĘCIE PROBLEMOWE POJĘCIA: Temat intencjonalności generuje pięć fundamentalnych pytań: 1) Czym jest relacja intencjonalna? 2) Czym są treści mentalne? 3) Czym są lub mogą być przedmioty intencjonalnego odniesienia aktów? 4) Intencjonalność w naukach społecznych: intencjonalność kolektywna. 5) Intencjonalność a *cognitive science*.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Do systematycznego badania intencjonalności w różnych obszarach przyczyniają się programy badawcze. W sekcji tej omówiono: 1) program

intencjonalności fenomenalnej, który dąży do głębszego zrozumienia intencjonalności w doświadczeniach sensorycznych, 2) program semantyzacji, który podkreśla silny związek języka i semantyki w kształtowaniu się myślenia i poznawania. Zaprezentowano również pojęcie intencjonalności kolektywnej, która przychodzi nam z pomocą w wyjaśnieniu natury ludzkiego życia społecznego. Ponadto pokazano, jak pojęcie intencjonalności może być intelektualnie i badawczo inspirujące dla badaczy z obszaru *cognitive science*.

Słowa kluczowe: Brentano, intencjonalność, intencjonalność mentalna, intencjonalność języka, intencjonalność kolektywna

Intencjonalność jest centralnym zagadnieniem filozofii umysłu – życie mentalne lub jednostkowa świadomość są intencjonalne. Stanowi też ważny problem filozofii języka, względnie semiotyki – znaki i wyrażenia językowe odnoszą się do czegoś lub na coś wskazują. Ale nie tylko. Problematyka intencjonalności podejmowana jest dzisiaj również w ramach filozofii społecznej i w kontekście nauk społecznych – obok „mojej intencjonalności” występuje intencjonalność kolektywna, „intencjonalność «my»”. Opracowane tutaj hasło porządkuje przede wszystkim zagadnienie intencjonalności w ramach tej pierwszej perspektywy (filozofia umysłu i języka). Ponieważ hasło napisane zostało na potrzeby serii Słowniki Społeczne, uwzględniono również perspektywę nauk społecznych i pojęcie intencjonalności kolektywnej.

Definicja pojęcia

Intencjonalność jest cechą naszych aktów mentalnych (takich jak np. pragnienie, przekonanie, wyobrażenie czy myślenie), świadomości czy umysłów generalnie i polega na ukierunkowaniu ich na pewien obiekt: pragnienie jest zawsze pragnieniem czegoś, przekonanie – przekonaniem o czymś, wyobrażenie – wyobrażeniem czegoś, myślenie – myśleniem o czymś itd. Można to krótko zamknąć w następującej definicji: „intencjonalność to ukierunkowanie aktów mentalnych na coś”. Ze względu na swoje językowe pochodzenie związana jest z intencją – pochodzi od łac. *intentio*, za pomocą którego scholastycy przełożyli gr. *noema* (pojęcie). Przy tym etymologia wskazuje na jedną z możliwych ścieżek interpretacyjnych tego pojęcia – na intencjonalność jako własność nakierowania umysłu na coś (lub identyfikację czegoś) za pomocą treści pojęciowej. W tym sensie intencjonalność związana jest ściśle z funkcjonowaniem umysłu i z naszą konceptualnością.

Od tak pojętej intencjonalności należy odróżnić intencjonalność kolektywną, która odnosi się do podmiotów społecznych i jest niezbędna do zrozumienia natury życia społecznego i instytucjonalnego. Mamy z nią do czynienia wtedy, gdy dwie osoby – lub więcej osób – dzielą wspólne intencje lub cele, a ich działania są koordynowane w sposób, który odzwierciedla to wspólne zaangażowanie. Nauki społeczne zajmują się badaniem społeczeństw (w tym przemian społecznych) oraz

interakcji międzyludzkich w różnych kontekstach: kulturowym, ekonomicznym czy politycznym. Badania intencjonalności zbiorowej to badania interdyscyplinarne, skoncentrowane na podstawach psychologicznych (m.in. na pragnieniach i potrzebach, intencjach, różnych formach myślenia i rozumowania) wspólnych lub współdzielonych działań.

Analiza historyczna pojęcia

Problem intencjonalnego odniesienia i mentalnej inegzystencji ma swoje źródła w starożytnej, a następnie średniowiecznej filozofii. W antyku pojawia się między innymi w filozofii Arystotelesa. Stagiryta w dziele *O duszy* podjął kwestię działania zmysłów i natury intelektu. Jego zdaniem każdy przedmiot w świecie jest złożony z materii i formy. Natomiast poznanie zmysłowe przedmiotów polega na przyjmowaniu form zmysłowych bez materii i na przekazywaniu ich do obróbki przez intelekt. Formy zmysłowe zawierają już w sobie formy umysłowe. Te ostatnie najpierw są magazynowane w intelekcie biernym, a następnie wykorzystywane przez intelekt czynny. Poznanie w filozofii Arystotelesa domaga się immanentnego pośrednika między rzeczą a myślą, którą stanowi forma. Anzelm przyjął rozróżnienie na istnienie w intelekcie (*in intellectu*) i istnienie w rzeczywistości (*in re*). To, co istnieje w rzeczywistości, jest obecny też w intelekcie. A jednak nie wszystko, co istnieje w intelekcie, musi istnieć w rzeczywistości. Ludzki umysł może bowiem generować abstrakcyjne idee, wyobrażenia czy koncepcje, które nie mają bezpośredniego odpowiednika w świecie zewnętrznym. Z kolei Tomasz z Akwinu „naucza, że to, o czym się myśli, istnieje intencjonalnie w podmiocie myślącym; przedmiot miłości – w kochającym; to, co pożądane, w pożądającym” (Brentano, 1999, s. 127). Dążąc do wyjaśnienia wielu kwestii związanych z działaniem umysłu teoretycznego i praktycznego (woli), Akwinata chętnie posługiwał się wieloznacznym terminem *intentio*, m.in. na określenie formy, pojęcia lub ponadzmysłowego wyobrażenia, czym antycypował niektóre wątki związane z intencjonalnością.

Problematyką intencjonalności w myśli średniowiecznej zajmował się systematycznie Dominik Perler, który zaproponował pięć modeli intencjonalności: 1) model formalnej identyczności Tomasza z Akwinu, 2) model intencjonalności konstytutywnej (odpowiedzialnej za przyrost

wiedzy) Piotra Oliviego i Dietricha von Freiberga, 3) model przedmiotów intencjonalnych Jana Dunsza Szkota, 4) model intencjonalnej obecności w umyśle Hervaeusa Natalisa i Piotra Auriola oraz 5) model znaków naturalnych Wilhelma Ockhama i Adama Wodehama (Perler, 2002). Współcześnie wieloma aspektami intencjonalności w epistemologii, metafizyce czy filozofii języka starożytnych i średniowiecznych myślicieli zajmują się m.in. Joel Biard, Myles Bumyeat, Victor Caston, Dorothea Frede, Richard Gaskin, Elizabeth Kruger, Cyrille Michon, Dominic O'Meara, Claude Panaccio, Robert Pasnau, Dominik Perler, Christof Rapp, Peter Simons, Richard Sorabji, Hermann Weidemann (Perler, 2001).

Do współczesnej filozofii problematyka intencjonalności została wprowadzona przez austriackiego psychologa i filozofa Franza Brentana, który próbował na nowo ustalić główny przedmiot psychologii opisowej oraz wypracować dla niej lepszą od dotychczasowej metodę badawczą. Za właściwą metodę Brentano uznał doświadczenie, ogląd i dokładny opis procesów psychicznych: „Podstawę psychologii, tak samo jak przyrodoznawstwa, stanowi spostrzeżenie i doświadczenie” (Brentano, 1999, s. 42). Za główny przedmiot psychologii przyjął natomiast akty mentalne, które charakteryzują się intencjonalnością, czyli nakierowaniem na pewien przedmiot. Podstawową charakterystykę zjawisk mentalnych znajdujemy w następującym fragmencie *Psychologii z empirycznego punktu widzenia*:

Każdy fenomen psychiczny charakteryzuje się tym, co średniowieczni scholastycy nazwali intencjonalną (lub też mentalną) [...] inegzystencją pewnego przedmiotu, a co my – jakkolwiek nie całkiem jednoznacznie – nazwalibyśmy odniesieniem do pewnej treści, skierowaniem na pewien obiekt (przez który nie należy tu rozumieć czegoś realnego) lub immanentną przedmiotowością. Każde [zjawisko psychiczne] zawiera coś jako obiekt, chociaż nie każde w ten sam sposób. W przedstawieniu coś jest przedstawiane, w sądzie uznawane lub odrzucone, w miłości kochane, w nienawiści nienawidzone, w pożądaniu pożądane itd. (Brentano, 1999, s. 126).

Z pierwszego zdania wynika, że intencjonalność to cecha, która wyraźnie odróżnia zjawiska psychiczne od zjawisk fizycznych – wszystkie akty mentalne, i tylko one, charakteryzują się intencjonalnością. Poza tym jednak cały ten fragment nie jest jasny i generuje wiele pytań: Czym są akty mentalne i jakiego rodzaju akty istnieją? Czym jest mentalna

inegzystencja? Czym jest skierowanie na pewien obiekt? Jakiego typu to relacja? Jakiego rodzaju są to obiekty? W jaki sposób obiekty te determinują relację odniesienia? Itd. Na pytania te padały różne odpowiedzi, co następnie skutkowało powstawaniem różnych ścieżek myślenia.

Ujęcie problemowe pojęcia

W punkcie wyjścia proponuję skoncentrować się na trzech problemach i rozważyć je krótko (częściowo uwzględniając późniejsze ustalenia): 1) Czym jest relacja intencjonalna? 2) Czym są treści mentalne? 3) Czym są lub mogą być przedmioty intencjonalnego odniesienia?

Ad 1) Intencjonalność to „nakierowanie na” lub „odnoszenie się do” jakiegoś przedmiotu, a zatem intencjonalność jest pewnego rodzaju relacją. Relacja bowiem w podstawowym sensie to stosunek jednego elementu do drugiego. Relacje mogą być różnego rodzaju. Jedne, fizyczne lub biologiczne, zachodzą między jakimiś obiektami lub organizmami, inne, personalne lub społeczne, zachodzą między osobami lub grupami społecznymi. Relacje te zakładają istnienie przedmiotów odnoszących się do siebie i determinują określone przedmioty lub zdarzenia, które mogą być zweryfikowane w czasoprzestrzeni – wszystkie człony tych relacji mają charakter czasoprzestrzenny. Innego rodzaju są natomiast relacje logiczne. Człony relacji logicznej to treści zdań logicznych, składniki zdaniowe lub inne wyrażenia logiczne, które są łączone za pomocą spójników logicznych. Najbardziej podstawowe relacje logiczne to koniunkcja, alternatywa, implikacja, równoważność i negacja. Relacje logiczne i obiekty logiczne są poza czasem i poza przestrzenią, są wieczne i niezmiennie. Człony relacji logicznych wciąż należą do tej samej kategorii. Natomiast podmiotowo-przedmiotowa relacja intencjonalna jest specyficzna i wyjątkowa. Jej wyjątkowość polega na tym, że drugi człon relacji intencjonalnej może być bardzo różny. Czasami jest to realny obiekt w czasoprzestrzeni lub tylko jego własności. Może to być zdarzenie, sytuacja lub stan rzeczy. Przedmiotem odniesienia mogą być również przekonania innych – pewien podmiot może być też przekonany, że inna osoba żywi takie lub inne przekonanie. Niektóre przedmioty tego odniesienia mogą w ogóle nie istnieć lub mogą być tylko pomyślane (obiekty abstrakcyjne) albo wyobrażone (obiekty

fikcjonalne). Z tego powodu uważa się, że relacja intencjonalna jest relacją wyjątkową, quasi-relacją lub pseudo-relacją.

Nasze umysły i wszystkie zjawiska psychiczne charakteryzują się intencjonalnością, której nie można przypisać temu, co fizyczne: żadne zjawisko fizyczne nie przejawia intencjonalności. Intencjonalność wymyka się zatem naukowemu wyjaśnieniu, bo wymyka się temu, co przyrodnicze. Nie da się jej wyjaśnić tak, jak wyjaśnia się świat przyrodniczy i relacje fizyczne. To zmobilizowało późniejszych filozofów do podejmowania różnych prób naturalizowania intencjonalności, treści czy znaczenia językowego. Do najbardziej znanych prób naturalizacji intencjonalności należy instrumentalizm Daniela C. Dennetta, teoretyczno-informatyczne podejście Freda Dretskego czy teleosemantyka Ruth Millikan.

Ad 2) Brentano odróżnił przedmiot transcendentny od immanentnego. Ten ostatni może być utożsamiany – choć nie musi – z treścią. Kazimierz Twardowski jednoznacznie opowiedział się za wyraźnym odróżnieniem aktu, treści i przedmiotu przedstawienia (Twardowski, 1965). Przez akt rozumiał jakąś czynność psychiczną, przez treść to, co sprawia, że każdy akt kieruje się na przedmiot. Natomiast przedmiot znajduje się poza aktem. Przy tym Twardowski akcentował niesamodzielną rolę aktu i treści (nie może istnieć akt bez treści i treść bez aktu) oraz samodzielność przedmiotu (przedmiot jest wyraźnie czymś innym niż akt i treść). Niezależności przedmiotu od aktu i samego aktu bronił Alexius Meinong. Treść mentalna odgrywa kluczową rolę w intencjonalności umysłu, ponieważ to dzięki niej podmiot może kierować się na coś, co istnieje w rzeczywistości zewnętrznej lub we własnym wnętrzu. Treść może być rozpatrywana nie tylko jako przedmiot psychologii czy filozofii umysłu, ale również jako przedmiot logiki i semiotyki – wówczas odnosi się ona do tego, co jest wyrażone przez sąd lub zdanie logiczne. Na razie aspekt ten pominiemy.

Istnieje zgoda co do tego, że treść mentalna to zawartość informacyjna percepcji, przekonań, pragnień, uczuć, myśli i innych stanów mentalnych w umyśle jednostki. Podmiot wykorzystuje je do interpretacji i identyfikacji obiektów oraz zdarzeń w świecie. W ogólności istnieje też zgoda co do wielorakich przyczyn treści. Jednym z głównych źródeł treści mentalnych są oczywiście doświadczenia sensoryczne i poznanie zmysłowe. Informacje o świecie zewnętrznym są przetwarzane przez

zmysły, a następnie reprezentowane w umyśle w postaci treści mentalnych. Drugim ważnym źródłem treści jest myślenie dyskursywne i refleksja, które mogą prowadzić do nowych ważnych wniosków i nowych idei, a zatem do nowych treści poznawczych. Ponadto istotną rolę w kształtowaniu treści mentalnych odgrywa język. Język bowiem kształtuje to, jak myślimy, jakie mamy przekonania, jak kategoryzujemy rzeczywistość i jak organizujemy swoje doświadczenia.

Dostrzegając te wielorakie źródła i różne aspekty treści, teoretycy zaproponowali kilka dodatkowych rozróżnień. Najbardziej znane i szeroko dyskutowane jest rozróżnienie na treść wąską (*narrow mental content*) i szeroką (*broad mental content*) oraz na treść konceptualną (*conceptual content*) i niekonceptualną (*nonconceptual content*). Treść wąska odnosi się do przedstawień, które zależą wyłącznie od wewnętrznych stanów umysłu jednostki i nie uwzględniają kontekstu zewnętrznego, nie odnoszą się do cech obiektów ani do relacji między nimi. Treść szeroka natomiast odnosi się do przedstawień, które uwzględniają kontekst zewnętrzny i relacje między umysłem a światem zewnętrznym. Krótko rzecz ujmując: treść szeroka jest bardziej bogata i złożona niż treść wąska.

Z kolei różnica między treścią konceptualną a niekonceptualną dotyczy sposobu, w jaki reprezentowane są informacje w umyśle. Treść konceptualna odnosi się do reprezentacji, które są oparte na pojęciach lub kategoriach poznawczych i używają języka lub innych systemów symbolicznych do przetwarzania informacji. Na przykład myśli dotyczące złożonych przedmiotów lub idei wyrażone za pomocą wyrażeń językowych są przykładem myśli o treściach konceptualnych. Treść niekonceptualna odnosi się do czegoś, co jest nam dane w sposób bardziej bezpośredni i obejmuje np. wrażenia sensoryczne, doświadczenia percepcyjne, np. takie jak widzenie kolorów (*qualia*), które nie są wyrażone za pomocą pojęć i są trudne do wyartykułowania w języku.

Ad 3) Przedmiot odniesienia intencjonalnego to kolejny ważny temat. Szczegółowe analizy na ten temat znajdujemy w fenomenologii Edmunda Husserla (Husserl, 1901) i Romana Ingardena (Ingarden, 1931). Ten drugi zaproponował interesującą koncepcję przedmiotu intencjonalnego zarówno jako korelatu, jak i wytworu aktu świadomości. Nie będziemy jej dalej analizować. Husserl w swojej filozofii wyszedł od określenia istoty świadomości: świadomość jest intencjonalna,

tzn. jest zawsze świadomością czegoś (*Bewusstsein-von...*). Intencjonalność świadomości nie jest jednak czymś prostym i jednorodnym, bo na świadomość składają się różne akty (spozreganie, wyobrażanie sobie, przypominanie, planowanie itd.), które mają własną specyfikę i odmienne przedmioty. Przedmiotem spostrzeżenia jest pozapodmiotowy przedmiot realny, przedmiotem wyobrażania sobie jest przedmiot wyobrażony (wyobrażenie), przedmiotem przypominania – wspomnienia, przedmiotem planowania – plany i przewidywania. Na ogólne pytanie, do jakiego rodzaju przedmiotów odnosimy się w świadomych aktach, odpowiedź nie może jednak brzmieć: a) odnosimy się do przedmiotów transcendentnych, b) odnosimy się do przedmiotów istniejących wewnątrz umysłu, c) odnosimy się do przedmiotów, które istnieją zarówno immanentnie, jak i transcendentnie. Wedle Husserla poprawna odpowiedź powinna brzmieć: w każdym z aktów odnosimy się do zewnętrznego przedmiotu, choć przedmioty te nie zawsze istnieją realnie. Intencjonalne myślenie jest bowiem niewrażliwe na istnienie lub nieistnienie pewnych przedmiotów. Odnoszenie się do czegoś to tylko jedna z wielu funkcji, jakie pełni intencjonalność w świadomym życiu podmiotu. Inne ważne funkcje to interpretacja, synteza i konstytucja. Ostatnia z wymienionych odpowiada za tzw. konstytucję świata w świadomości. Choć kwestia intencjonalności była czymś fundamentalnym w fenomenologii Husserla, to nigdzie nie otrzymujemy jasnego jej wykładu czy wyraźnej teorii. Wraz z kolejnymi analizami, zwłaszcza z perspektywy idealizmu transcendentnego, Husserl odkrywa jej kolejne aspekty i funkcje. Intencjonalność nie jest zwykłą cechą odróżniającą to, co mentalne, od tego, co fizyczne, ale jest w naszej świadomości czymś dynamicznym i twórczym (Husserl, 1975).

Natomiast wspomniany już Twardowski naturę przedmiotu próbuje uchwycić niezależnie od samego aktu i treści przedstawienia. Rozumie go stosunkowo szeroko. Każde przedstawienie ma jakiś przedmiot – nie istnieją przedstawienia bezprzedmiotowe. Do istoty przedmiotu należy jego przedstawialność. Może on być albo realny albo nierealny, możliwy lub niemożliwy, istniejący lub nieistniejący. Analizami jego zainspirował się Alexius Meinong, dla którego przedmiotem jest wszystko, co można uchwycić, rozpoznać i ocenić za pomocą aktu umysłowego (Meinong, 1904). Przedmiotem tym może być wszystko, co możliwe i niemożliwe. W konsekwencji Meinong zaproponował własną teorię przedmiotu

(*Gegenstandstheorie*), w ramach której wyróżnił następujące typy przedmiotów: 1) realne (jako istniejące), 2) idealne (w terminologii Meinonga jako subsystujące), 3) nieistniejące (sprzeczne, np. kwadratowe koło lub niesprzeczne, np. obecny król Polski), 4) bytowo nieokreślone (niepełne, jak m.in. przedmioty fikcyjne, np. Hamlet). Należy przy tym pokreślić, że 3) i 4) nie wykluczają się.

Teoria przedmiotu (w tym przedmiotu nieistniejącego) Meinonga z jednej strony spotkała się z krytyką ze strony Bertranda Russella, a z drugiej przyczyniła się do rozwoju innych teorii przedmiotu nieistniejącego (Dale Jacques, Ernst Mally, Terence Parsons, Richard Routley, znany również jako Richard Sylvan, Edward N. Zalta). Inni obrali różne strategie redukcjonistyczne: 1) strategia deontologizacji sprowadza byty do poziomu języka i zdań: nie ma nieistniejących obiektów, ale istnieją prawdziwe zdania o nieistnieniu obiektów (Tim Crane, Charles Crittenden); 2) strategia światów możliwych: nie ma nieistniejących przedmiotów w aktualnym świecie, ale mogą istnieć w jednym z możliwych światów (Francesco Berto, Graham Priest); 3) fikcjonalizm: choć przedmioty nieistniejące w rzeczywistości nie istnieją, to jednak posiadają pewien rodzaj istnienia fikcyjnego i pełnią swoją funkcję w ramach fikcyjnego świata, w którym istnieją (Mark Kalderon, Frederick Kroon, Kendall Walton, Stephen Yablo). Na gruncie polskim problematyką przedmiotów zajmowali się m.in. Arkadiusz Chrudzinski, Jacek Paśniczek, Maciej Sendłak.

Refleksja systematyczna z wnioskami i rekomendacjami

Prymat indywidualnej intencjonalności mentalnej. Być może najbardziej znaną koncepcją intencjonalności świadomości jest koncepcja Johna Searle'a. Filozof traktuje świadomość jako funkcję biologiczną, która ewoluowała w procesie naturalnej selekcji, ale jest ona również wynikiem specyficznej organizacji mózgu. Świadomość to rezultat działania neurofizjologicznych procesów zachodzących w mózgu. Równocześnie Searle podkreśla, że samo poznanie tych procesów nie wyjaśnia natury świadomości. Świadomość jest subiektywnym doświadczeniem każdego jednostkowego umysłu. To znaczy, że tylko osoba doświadczająca

świadomości może ją zrozumieć, a jej doświadczenie jest dostępne jedynie dla niej samej. Wskazuje również na jakościowy charakter subiektywnych doświadczeń, które nie są łatwe do skonceptualizowania i wyartykułowania (*qualia*). Tworzą go m.in. doświadczenie widoku tęczy, doświadczenie zapachu czerwonej róży, doświadczenie bólu itp. Poza tym ważną cechą świadomości jest intencjonalność: świadomość zawsze odnosi się do czegoś, zawsze czegoś dotyczy (Searle, 1983).

Odwołując się do koncepcji intencjonalności Searle'a, warto zwrócić uwagę przynajmniej na cztery sprawy. Po pierwsze, jego zdaniem nie wszystkie świadome stany mentalne są intencjonalne – istnieją nie-intencjonalne doświadczenia (radosne uniesienie, przygnębienie czy melancholia), a ponadto istnieją nieświadome stany intencjonalne, ponieważ istnieją nieświadome spostrzeżenia, przekonania i myśli. Po drugie, intencjonalne są przede wszystkim różne stany mentalne, które tworzą świadomość. Tylko świadomości i tworzącym ją stanom przysługuje intencjonalność wewnętrzna (pierwotna). Poza tym intencjonalne są język i inne wytwory ludzkiej twórczości, ale im przysługuje intencjonalność wtórna. Po trzecie, poszczególne stany mentalne składają się z treści intencjonalnej (stany mają określoną zawartość, która reprezentuje obiekty, zdarzenia, stany rzeczy itp.) i trybu psychologicznego, który wpływa na sposób, w jaki treść intencjonalna jest postrzegana i doświadczana przez podmiot. Po czwarte, intencjonalne stany mentalne charakteryzują się kierunkiem dopasowania („umysł – świat”, „świat – umysł”) i posiadają warunki fortunności: 1) prawdziwości (przekonanie) – aby zdanie intencjonalne, takie jak „John jest przekonany, że słońce świeci”, było uznane za prawdziwe, musi istnieć rzeczywisty stan rzeczy, dzięki któremu John faktycznie jest przekonany, że jest słonecznie; 2) spełnienia (pragnienia) – jeżeli John ma pragnienie, na przykład chciałby się napić herbaty, warunkiem spełnienia tego pragnienia jest dostępność herbaty w danym momencie i możliwość jej konsumpcji; 3) realizacji (zamiaru) – jeżeli John ma zamiar zrobić coś konkretnego, na przykład pójść na spacer do parku Jordana, warunkiem realizacji tego zamiaru jest podjęcie działań związanych z pójściem na spacer do parku Jordana, który rzeczywiście istnieje.

Często wskazuje się na dwie podstawowe cechy świadomości: 1) jakościowe doznanie lub subiektywne doświadczenie (np. doświadczenie kolorów tęczy, zapachu róży, swędzenia lub bólu) oraz

2) intencjonalność. Przynajmniej niektórzy badacze świadomości uważają, że wszystkie stany świadome są albo stanami fenomenalnymi, albo intencjonalnymi i odpowiednio do tego przyjmują dwa aspekty lub dwa rodzaje świadomości (np. Ned Block). Jednak ze względu na pewne kontrprzykłady (żywe i emocjonalne myślenie, celem którego jest wyjście z trudnej sytuacji, jest nie tylko intencjonalne, ale ma również charakter fenomenalny), trudno zgodzić się na to, że istnieją dwa różne rodzaje świadomości. W tym kontekście bardziej sensowne wydaje się mówienie o dwóch aspektach świadomości. Tak czy inaczej, relacje między świadomością i intencjonalnością, świadomością i subiektywnym doświadczeniem oraz intencjonalnością i subiektywnym doświadczeniem wcale nie są oczywiste i wymagają ponownego przemyślenia – zwłaszcza w kontekście ostatnich badań nad intencjonalnością świadomego doświadczenia. Mam na myśli przede wszystkim program intencjonalności fenomenalnej (*phenomenal intentionality program*).

Program ten powstał w celu zgłębienia natury intencjonalności doświadczenia fenomenalnego. Podstawowym jego założeniem jest to, że wszelkie doświadczenie sensoryczne jest intencjonalne. Zwolennicy tego programu bronią przynajmniej trzech poglądów. Po pierwsze, bronią prymatu fenomenologii, a dokładniej: prymatu doświadczenia w wyjaśnianiu, w tym wyjaśnianiu intencjonalności. Doświadczenie zmysłowe jest w dużym stopniu intencjonalne, a jego intencjonalność jest zakorzeniona w jego fenomenologii. Po drugie, bronią internalizmu fenomenalnego i odrzucają eksternalizm. Intencjonalność fenomenalna jest zdeterminowana wewnątrznie, co oznacza, że nie zależy od czynników zewnętrznych. Po trzecie, bronią liberalizmu fenomenalnego. Oprócz fenomenologii zmysłowej istnieje również fenomenologia poznawcza, która pomaga ugruntować treść stanów poznawczych. To podejście ma rozwiązać problemy dotyczące determinacji treści w kontekście filozofii języka i poznania. Wygląda na to, że program intencjonalności fenomenalnej próbuje dokonać poważnej rewolucji, jeśli nie w filozofii umysłu w ogóle, to przynajmniej w myśleniu o intencjonalności i świadomym doświadczeniu (Kreigel, 2013; Pautz, 2013).

Prymat semantyki języka. Wielu teoretyków intencjonalności, w tym już wspomniani (Brentano, Twardowski, Meinong, Husserl, Ingarden, Searle), dostrzegali działanie intencjonalności w języku i rozważali m.in.

kwestie znaczenia i odniesienia wyrażen językowych. Intencjonalność w języku przejawia się bowiem przede wszystkim w relacji języka do rzeczywistości pozajęzykowej, w odniesieniu nazw do przedmiotów i zdań do stanów rzeczy lub faktów. Niektórzy dostrzegali słabość mentalistycznej perspektywy i krytykowali „introspekcję”, „podmiot”, „subiektywność”, „akty”, „treść mentalną” i samą ideę intencjonalności mentalnej. Czynili to mimochodem, fragmentarycznie lub programowo. Inni nie oglądali się na mentalistów i od razu zwracali się w stronę logiki, semantyki i filozofii języka, czyli tam, gdzie była większa szansa na większą precyzję, niezawodność i obiektywność. Na szczególną uwagę zasługuje semantyka Gotloba Fregego, która – obok rozwiązań Bertranda Russella – zainspirowała wielu filozofów do badania języka i intencjonalnego myślenia poprzez język.

Frege dokonał ważnego spostrzeżenia na temat identyczności sądów „ $a = a$ ” i „ $a = b$ ”. Sąd „Gwiazda Wieczorna = Gwiazda Wieczorna” jest tautologią, ale zdanie „Gwiazda Wieczorna = Gwiazda Poranna” już nie jest tautologią. Nazwy tu podane, chociaż odnoszą się do tego samego przedmiotu (planeta Wenus), różnią się swoim sensem – sposobem dania przedmiotu. W rezultacie Frege przyjmuje, że znaki językowe nie tylko coś oznaczają, mają swoje odniesienie (*Bedeutung*), ale posiadają również pewien sens (*Sinn*). A mówiąc precyzyjniej, nazwy denotują przedmiot i wyrażają sens – sens nazwy determinuje przedmiot odniesienia. Oprócz nazw Frege uwzględnia predykaty, które oznaczają wyłącznie pojęcia (idea, czysto myślowe przedstawienia sobie czegoś) oraz zdania. Zdanie wyraża pewną logiczną myśl i denotuje wartość logiczną prawdy lub fałszu; myśl determinuje wartość logiczną. Ponadto w ramach swojej semantyki Frege przyjmuje kilka zasad. 1) *Zasada zastępowalności*: jeśli dwa wyrażenia mają ten sam sens, to mogą być wzajemnie zamienione w kontekście dowolnego zdania, nie zmieniając jego wartości logicznej. 2) *Zasada kompozycjonalności*: znaczenie (sens) wyrażenia złożonego zdeterminowane jest w całości poprzez znaczenie części składowych oraz reguł ich łączenia – znaczenie takiego wyrażenia może być zrozumiane tylko wtedy, gdy rozumie się jego części składowe. 3) *Zasada kontekstualności*: znaczenie wyrażenia zależy od kontekstu, w którym jest używane. Słowa mają różne znaczenia, a zatem o znaczenie słowa należy pytać zawsze w kontekście całego zdania.

Prace Fregego stanowiły fundament dla dalszych badań nad semantyką języka. Jego rozróżnienie między sensem i odniesieniem, zasada kompozycyjności czy w ogóle analiza struktury językowej stanowiły podstawę dla wielu współczesnych teorii semantycznych. Bertrand Russell podjął jego idee w pracy nad logiką i semantyką języka, poddał krytyce rozróżnienie na sens i odniesienie i rozwinął teorię denotacji i deskrypcji. Ludwig Wittgenstein inspirował się Fregem, pisząc *Traktat logiczno-filozoficzny* (w większości dyskutując z Fregem, poważnie modyfikując jego pojęcia i idee). Rudolf Carnap pracował nad semantyką języka formalnego, a Donald Davidson nad semantyką języka naturalnego, teorią znaczenia i prawdy. Dokonania niemieckiego logika miały wpływ również na filozofię umysłu. Jego analizy semantyczne wniosły wkład w zrozumienie natury myślenia, konceptualizacji i reprezentacji mentalnych.

Do kluczowych zagadnień w filozofii języka i semantyce należą: problem znaczenia językowego oraz referencji. Zdania języka naturalnego mają znaczenie i na jego mocy – podobnie jak stany umysłu – ukierunkowane są na rzeczy inne niż one same. Ze znaczeniem wiążą się też inne pytania, np. o to, czym są reguły i konwencje, które nadają znaczenie poszczególnym wyrażeniom, oraz jakie istnieją zasady interpretacji tych wyrażeń przez odbiorców. Tutaj problem znaczenia pominiemy i skoncentrujemy się na referencji, czyli na problemie odniesienia językowego. Analiza tego, w jaki sposób terminy odnoszą się do swoich desygnatów, pomaga nam zrozumieć strukturę i funkcjonowanie języka, ale jest również gwarantem lepszego opisu i poznania świata, jak też efektywnej komunikacji między ludźmi. Istnieją różne teorie dotyczące mechanizmu odniesienia terminów referencyjnych, w zależności od tego, jakiego rodzaju terminy chcemy uwzględnić. Niektóre teorie sugerują, że odniesienie jest możliwe poprzez zrozumienie znaczenia terminu, inne podkreślają rolę kontekstu czy środowiska w procesie odniesienia, a jeszcze inne zakładają bardziej złożone mechanizmy, takie jak teoria odniesienia oparta na intencjach. Szczególnie interesującym przypadkiem odniesienia jest odniesienie bezpośrednie.

Teoria bezpośredniego odniesienia, jak wskazuje David Kaplan, sugeruje, że funkcja pewnych wyrażeń językowych, takich jak nazwy własne, indeksy czy zaimki wskazujące, polega na bezpośrednim wskazywaniu na konkretne obiekty w świecie, a nie dokonywaniu tego poprzez

pośrednie opisy czy cechy tych obiektów. Na przykład, gdy używamy nazwy własnej „Warszawa”, nie odnosimy się poprzez opis, lecz odnosimy się bezpośrednio do miasta, które nosi tę nazwę. Na idei bezpośredniego odniesienia opierają się przyczynowe teorie odniesienia Saula Kripkego, Keitha Donnellana czy Hilarego Putnama, które powstały jako odpowiedź na słabości deskrypcyjnego podejścia Russella.

W książce *Nazywanie a konieczność* Kripke twierdzi np., że nazwy własne, w przeciwieństwie do większości opisów, są sztywnymi desygnatami, czyli wyrażeniami językowymi, które niezmiennie oznaczają ten sam obiekt we wszystkich możliwych światach (Kripke, 1988). Dzieje się tak na mocy chrztu udzielonego pewnej osobie, np. prezydentowi USA R. Nixonowi. Sztywność dotyczy nie tylko nazw własnych, ale również nazw gatunkowych odnoszących się do rodzajów naturalnych („tygrys”, „złoto” czy „woda”). Obejmują one obiekty, które posiadają pewne cechy esencjalne, konieczne (niezależnie od tego, że mogą mieć też przypadkowe). Przykładowo, jeśli mówimy o „wodzie” jako rodzaju naturalnym, to cechą esencjalną tej wody jest skład chemiczny H₂O. Chociaż woda może być w różnym stanie i może mieć różne ilości zanieczyszczeń, to wciąż pozostaje wodą, jeżeli tylko posiada esencjalną cechę chemicznego składu H₂O.

Jak widać, idea intencjonalności jako pewnego rodzaju odniesienia inspirowała wielu filozofów, którzy uprzywilejowali analizę języka. Na koniec warto może wspomnieć jeszcze o propozycjach Wilfrida Sellarsa i Donald Davidsona, których koncepcje wpisują się w program semantykalizmu. Program ten przyjmuje jako podstawową tezę o pierwotnej relacji semantycznej i próbuje wyjaśnić intencjonalność świadomości (i to wszystko, co ją tworzy) za pomocą własności semantycznych języka. Sellars był zwolennikiem tezy, że semantyka języka i intencjonalność myślenia dyskursywnego są ze sobą ściśle powiązane. Uważał, że społecznie ukształtowany język odgrywa kluczową rolę w kształtowaniu się „języka mentalnego”, a co za tym idzie – język ten ma wpływ na nasz sposób poznawania i rozumienia świata (Sellars, 1956).

Stanowisko Davidsona jest jeszcze silniejsze. Twierdzi on, że język jest warunkiem koniecznym myślenia i posiadania przekonań – nie można myśleć i poznawać bez języka. Jego argumentacja będzie mniej więcej tak: aby mieć przekonania i myśli, musimy być w stanie interpretować

język. Interpretowanie języka oznacza rozumienie treści wypowiedzi innych osób oraz przypisywanie znaczenia wyrażeniom językowym. Interpretujemy za pomocą języka, a zatem język jest czymś podstawowym i kluczowym. Bez języka niemożliwe są myślenie i poznawanie; zwierzęta nie mają języka, a zatem nie potrafią myśleć i poznawać. Jedynie ludzie dysponują językiem, co implikuje, że tylko ludzie są zdolni do myślenia i poznawania (Davidson, 1982).

Intencjonalność kolektywna. Ludzie są istotami społecznymi: posługują się językiem i angażują się w interakcje społeczne z innymi członkami pewnej grupy w ramach swojego gatunku. Rozwijają myślenie kooperatywne, nawiązują współpracę i zrzeszają się w grupy. Dla ludzi wspólnota jest kluczowym kontekstem, w którym tworzą oni normy społeczne, przestrzegają ich i żyją według wspólnych wartości. Życie społeczne człowieka przyczynia się do rozwoju życia instytucjonalnego i kulturalnego oraz do życia w ramach struktur politycznych, w których podejmowane są decyzje dotyczące wspólnego życia. Polityka to nic innego jak zdolność zarządzania wspólnotą (państwem) dla dobra wspólnego. Te społeczne, kulturowe i polityczne aspekty ludzkiego funkcjonowania mają swoje źródło w intencjonalności kolektywnej.

Podwaliny pod współczesną filozoficzną dyskusję na temat intencjonalności kolektywnej położyli filozofowie starożytni, a później nowożytni (Platon, Arystoteles, Hobbes, Locke, Rousseau), jak i socjologowie (Durkheim, Simmel, Weber). Wielu współczesnych teoretyków wskazuje również na inspirujące myśli W. Sellarsa (Tuomela & Miller, 1988). Poważny wkład w rozwój badań nad intencjonalnością kolektywną wnieśli m.in. Margaret Gilbert, John Searle i Raimo Tuomela. Ta pierwsza postawiła pytanie o znaczenie słowa „my”. Zbadła szereg przypadków, w których użycie tego wyrażenia jest niewłaściwe, by w końcu wskazać na centralne znaczenie słowa „my” – takie, w którym „my” odnosi się do podmiotu w liczbie mnogiej, do podmiotu kolektywnego (*plural subject*), którego nie da się zredukować do własności poszczególnych podmiotów. John Searle, korzystając z kategorii intencjonalności mentalnej, zbadał mechanizmy ludzkiego umysłu, a następnie wprowadził koncepcję intencjonalności kolektywnej oraz teorię funkcji statusowych i reguł konstytutywnych. Dzięki tym narzędziom wyjaśnił, w jaki sposób rzeczywistość instytucjonalna i struktury społeczne są konstruowane i utrzymywane przez ludzkie uzgodnienia i działania. Natomiast Raimo Tuomela,

fiński filozof społeczny, w wyjątkowo wyrafinowany sposób pokazuje, jak człowiek balansuje w życiu między trybem ja (*I-mode*) i trybem my (*we-mode*), a przede wszystkim jak funkcjonuje w perspektywie „my”, która obejmuje kolektywną intencjonalność. Intencjonalność ta jest głęboko wbudowana w myślenie i działanie istot ludzkich i jest rezultatem adaptacji zarówno biologicznej, jak i kulturowej.

Intencjonalność kolektywna jest intencjonalnością wspólną pewnej grupie lub też intencjonalnością „my” i przejawia się w tym, że osoby należące do tej grupy mają wspólne przekonania, a zwłaszcza intencje, podejmują wspólne zadania lub wybierają i zmierają w tym samym celu. Intencjonalność indywidualna wyraża się w zdaniach: „widzę, że x”, „sądzę, że y”, „mam nadzieję, że z”, „zamierzam osiągnąć to a to”, natomiast intencjonalność kolektywna w zdaniach: „widzimy, że x”, „sądzimy, że y”, „mamy nadzieję, że z”, „zamierzamy osiągnąć to a to”. Intencjonalność kolektywna leży u podstaw gry piłkarskiej pewnej drużyny, której celem jest wygrana w meczu. Każdy gracz w drużynie może mieć swoje indywidualne cele w czasie treningu, a później indywidualne zadania w trakcie meczu, ale ostatecznie wszyscy jako drużyna dążą do osiągnięcia wspólnego celu – do wygranej. Inny przykład: działanie zespołu nad wspólnym projektem. Każdy członek zespołu może mieć swoje indywidualne zadania do wykonania, ale wszyscy dążą do zrealizowania wspólnego celu projektu. Badania na tym polu pomagają nam zrozumieć, jak grupy wyznaczają cele, podejmują decyzje, organizują działania i – ostatecznie – osiągają zamierzone cele.

Koncepcja intencjonalności kolektywnej głosi, że grupy społeczne mogą mieć własne zamiary i cele, które wykraczają poza sumę intencji i celów poszczególnych jednostek składających się na daną grupę. Oznacza to, że grupa może działać jako odrębny podmiot z własnymi interesami i celami, które niekoniecznie są odzwierciedleniem intencji i celów jej poszczególnych członków. Na marginesie warto może podkreślić, że źródeł tego rodzaju myślenia można szukać już u Emila Durkheima, wedle którego rzeczywistość społeczna jest dla jednostki rzeczywistością zastaną, która w dużym stopniu determinuje życie jednostki. Społeczeństwo wywiera wpływ na swoich członków choćby poprzez język czy systemy norm i instytucji, które kształtują zachowania, oczekiwania i możliwości. Durkheim uważał, że indywidualne decyzje są często wynikiem społecznych nacisków, a nie tylko osobistych

preferencji. Społeczeństwo jako całość tworzy ramy, w których jednostki od początku jakoś funkcjonują, a to świadczy o jego niezależności.

Nie wszyscy badacze intencjonalności kolektywnej podzielają pogląd, zgodnie z którym grupy społeczne mają zamiary i cele zupełnie niezależne od tych, jakie mają jednostki. Przykładowo Michael Bratman uważa, że indywidualne zamiary są również ważne i niezbędne do wyjaśnienia tego, co wspólne. W swojej definicji uwzględnia trzy kluczowe elementy, niezbędne do istnienia współdzielonej intencji: 1) Zarówno ja, jak i ty, jako uczestnicy współdzielonej intencjonalności, musimy mieć własne intencje (intencje indywidualne), dotyczące realizacji danego celu czy działania. 2) Zarówno ja, jak i ty musimy mieć intencje, aby działać jako grupa i realizować wspólny cel; te intencje powinny być zgodne i harmonizować ze sobą (nie mogą być sprzeczne). 3) Zarówno ja, jak i ty musimy mieć wspólne zrozumienie naszych intencji – każdy z nas powinien wiedzieć, że ja mam intencję, ty masz intencję i że te intencje są wzajemnie przez nas uznawane.

Skąd bierze się zdolność do przyjmowania wspólnych intencji? Częściowej odpowiedzi na to pytanie dostarcza ewolucja (przy założeniu, że wspólna intencjonalność jest warunkiem wstępnym społeczeństwa kooperatywnego). Wiele zwierząt prowadzi samotniczy tryb życia (poza okresem godowym): niedźwiedzie polarne, lamparty, nosorożce, koala czy leniwce. Zwierzęta te lubią samotnie wędrować i polować. Inaczej małpy człekokształtne – te należą już do zwierząt wysoko uspołecznionych. Interesują się sobą nawzajem, porozumiewają się ze sobą (za pomocą takich sygnałów jak dźwięki, gesty, mimika), dokonują rytualnej pielęgnacji. Wspólnie polują na inne zwierzęta. W grupie tworzą hierarchię, której przewodzi samiec alfa. Choć zachowują wyraźne granice pomiędzy poszczególnymi grupami, potrafią względnie pokojowo współegzystować, a nawet korzystać z tego samego źródła pożywienia. Jednak nie spędzają ze sobą czasu i nie kooperują ściśle ze sobą, a często zachowują się też wrogo względem siebie. Dużo bardziej przyjazne są szympansy bonobo. Tutaj przedstawiciele różnych grup wchodzi już w ściśle interakcje, przebywają razem, dzielą się posiłkami. Postępują tak samo wobec przedstawicieli odrębnej grupy, jak wobec własnej. Człowiek jest najbliższym spokrewnionym z szympancami, przejął ich społeczny charakter, a nawet znacznie je w tym przewyższył. Dziś tworzy złożony świat kultury i struktury społeczne, które są często

oparte na hierarchiach społecznych, normach i złożonych systemach współpracy. Ewolucja biologiczna nie jest w stanie wyjaśnić tego skoku, ale dokonać tego może ewolucja kulturowa.

Rozwój człowieka jest głęboko zakorzeniony w społeczno-kulturowej rzeczywistości, która sama ewoluuje i kształtuje ludzkie procesy poznawcze, komunikację i twórczość. Kwestie te szczegółowo badał m.in. Michael Tomasello. Uczony podkreśla, że tylko ludzie są zdolni do tego, aby 1) konceptualizować jedną i tę samą sytuację z różnych perspektyw społecznych, 2) wyciągać społecznie rekurencyjne, autorefleksyjne wnioski na temat stanów intencjonalnych innych osób, 3) kontrolować i oceniać własne myśli w odniesieniu do normatywnych perspektyw i standardów innych grup. Ten wyraźnie inny sposób myślenia nazywa Tomasello obiektywno-refleksyjno-normatywnym. Do jego powstania przyczyniła się ludzka zdolność do współpracy i do dzielenia się intencjami, co nazywa Tomasello „współdzieloną intencjonalnością”. Różne formy współpracy wymagały od ludzi zdolności rozpoznawania celów i planów działania innych, co z kolei stymulowało rozwój języka, myślenia abstrakcyjnego i bardziej złożonych form komunikacji. To z kolei umożliwiło ludziom planowanie przyszłości, organizację większych grup społecznych oraz tworzenie bardziej złożonych systemów normatywnych (Tomasello, 2014).

A zatem dla pełniejszego zrozumienia rozwoju intelektualnego i społecznego człowieka kluczowe jest połączenie perspektyw ewolucji biologicznej i ewolucji kulturowej. Ewolucja biologiczna dostarcza podstawowych warunków dla powstania pierwotnych form wspólnego działania, które są niezbędne dla początkowego etapu rozwoju ludzkich zdolności poznawczych i społecznych. Jednakże równie istotna jest ewolucja kulturowa. Interakcje między biologią a kulturą generują złożone formy myślenia, komunikacji oraz zachowań społecznych, które definiują ludzką specyfikę. Wzajemne oddziaływanie tych dwóch sfer ewolucji pozwala na osiągnięcie poziomu zaawansowania, który niemożliwy byłby przy działaniu tylko jednej z nich.

Analiza intencjonalności kolektywnej przyczynia się do lepszego zrozumienia następujących kwestii: 1) na czym polega współdziałanie w grupach; 2) jaka jest natura wspólnego podejmowania decyzji; 3) na czym polega odpowiedzialność zbiorowa; 4) jak powstają i ewoluują struktury społeczne; 5) jaki jest ich status ontologiczny? (Lista ta nie

jest wyczerpująca i zawiera jedynie przykładowe zagadnienia). Nie ma miejsca tutaj na rozwinięcie tych kwestii. Odnotujmy jednak krótko, o co chodzi w każdej z wymienionych kwestii.

Ad 1) Kolektywne działanie to działanie podejmowane przez grupę, która ukierunkowana jest na osiągnięcie wspólnego celu, przy wspólnej akceptacji środków i aksjologii. Zakłada m.in. wspólne planowanie, komunikację i efektywne współdziałanie, które polega na dzieleniu się zadaniami i wzajemnym wsparciu w osiąganiu obranego celu. Kolektywne działanie wymaga zaufania. Gdy członkowie pewnej grupy ufają sobie nawzajem, komunikacja staje się łatwiejsza, a współpraca bardziej płynna i efektywna.

Ad 2) Kwestia budżetu domowego, wychowania dzieci czy planowanie różnych wydarzeń w rodzinie wymagają wspólnych decyzji. Decyzje tego rodzaju podejmowane są również w instytucjach publicznych, organizacjach rządowych i pozarządowych. Kolektywne podejmowanie decyzji wymaga otwartej komunikacji i negocjacji, w których członkowie grupy starają się osiągnąć porozumienie lub zdrowy kompromis. Negocjacje mogą dotyczyć takich aspektów decyzji jak priorytety, zasoby czy konsekwencje. Indywidualne decyzje są stosunkowo proste i szybko zapadają, ale charakteryzują się ograniczonym dostępem do informacji. Decyzje kolektywne są trudniejsze i bardziej czasochłonne, ale umożliwiają bardziej wszechstronną i kompleksową ocenę sytuacji. Nie znaczy to, że są wolne od błędów.

Ad 3) Członkowie pewnej grupy, podejmując wspólnie decyzje i realizując określone cele, ponoszą wspólną odpowiedzialność, nawet jeśli nie wszyscy bezpośrednio przyczyniają się do tych decyzji, działań i ich rezultatów. Członkowie grupy uczestniczą w profitach, jakie wynikają ze wspólnego działania, ale mogą też być moralnie i prawnie zobowiązani do podejmowania działań, które pomogą zrekompensować szkody w przypadku błędnych decyzji i szkodliwych działań.

Ad 4) Struktury społeczne mają swój początek w prostych interakcjach międzyludzkich, które z czasem utrwalają się w powtarzalne wzorce. Gdy pewne działania są regularnie powtarzane i uznane za normatywne, mogą stać się instytucjami. Tak powstałe społeczności nieustannie ewoluują i dostosowują się do zmieniających się warunków, takich jak: zmiany demograficzne, konflikty społeczne, katastrofy naturalne czy rozwój nowych technologii.

Ad 5) W kwestii statusu ontologicznego dominują trzy stanowiska. Nominaliści uważają, że struktury społeczne są jedynie produktami ludzkich działań i interakcji, a ich „realność” jest wtórna wobec jednostek. Dla konstruktywistów są one wynikiem społecznego konstruowania rzeczywistości. Realiści społeczni twierdzą natomiast, że struktury społeczne mają własną naturę i istnieją niezależnie od jednostek. Ich zdaniem instytucje społeczne przetrwają nawet po wymianie wszystkich swoich członków.

Intencjonalność a cognitive science. Kognitywistyka to interdyscyplinarna dziedzina badań, której przedmiotem jest umysł (względnie: system poznawczy), jego powstanie, struktura i działanie. Składają się na nią filozofia (zwłaszcza filozofia umysłu i epistemologia), psychologia (zwłaszcza psychologia poznawcza), neurobiologia, językoznawstwo kognitywne, sztuczna inteligencja i antropologia. Przedstawiciele tych dyscyplin wykorzystują różne naukowe i eksperymentalne metody, aby wyjaśnić rozmaite aspekty działania umysłu, a następnie współpracując ze sobą, starają się zintegrować zdobytą wiedzę. Wyjaśnianie umysłu z pomocą reprezentacji jest jednym z kluczowych podejść w dziedzinie kognitywistyki (mimo silnej opozycji ze strony antyrepresentacjonalistów). Te zaś ściśle powiązane są z intencjonalnością. Intencjonalność umysłu wyraża się bowiem poprzez reprezentacje umysłowe, które są podstawowymi jednostkami przetwarzania informacji w umyśle, analogicznymi do symboli lub obrazów mentalnych, które odzwierciedlają obiekty i zdarzenia w świecie zewnętrznym, ale mogą też reprezentować coś, czego w świecie zewnętrznym nie ma.

A zatem, po pierwsze, intencjonalność umożliwia badanie sposobu, w jaki umysł reprezentuje i manipuluje informacjami w celu osiągnięcia określonych rezultatów poznawczych, np. dochodzenia do prawdy czy rozwiązywania problemów. Po drugie, kognitywiści mogą modelować różne procesy poznawcze (pragnienia, przekonania, intencje), analizując, jak intencjonalność wpływa na planowanie działań i podejmowanie decyzji. Widzieliśmy powyżej, że intencjonalność odgrywa kluczową rolę w semantyce języka naturalnego. Zatem, po trzecie, kognitywiści mogą badać, jak intencjonalność wpływa na procesy interpretacji i komunikacji oraz rozumienia i generowania wypowiedzi językowych. Systemy poznawcze poznają i komunikują się, żeby działać. Kognitywiści, to po czwarte, mogą badać, jak intencjonalność kształtuje nasze

interakcje społeczne. Zrozumienie intencjonalności może też przyczynić się do rozwoju sztucznej inteligencji poprzez modelowanie umysłowych procesów i zachowań na podstawie intencji i rozumowań praktycznych. Kognitywiści zorientowani na sztuczną inteligencję lub robotykę mogą wykorzystać tę wiedzę do projektowania inteligentnych agentów i systemów komputerowych, które wykazują zdolność do rozumienia i interpretacji ludzkich intencji.

* * *

Dyskusja na temat intencjonalności toczy się dzisiaj w wielu dziedzinach i dyscyplinach, od filozofii umysłu po nauki społeczne i międzydziedzinowe (*cognitive science*), co pokazuje jej złożoność, bogactwo intelektualne i interdyscyplinarny charakter. Z punktu widzenia filozoficznego warto może podkreślić, że samo przyjęcie intencjonalności nie musi z góry determinować konkretnego stanowiska ontologicznego, epistemologicznego czy semantycznego. Oto dwa interesujące przykłady: analiza fenomenologiczna intencjonalności świadomości Husserla doprowadziła go do idealizmu, natomiast realizm fenomenologiczny Ingardena do przyjęcia bogatej ontologii, która broni bytów realnych, abstrakcyjnych i idealnych. Ostatecznie analiza intencjonalności otwiera drzwi do różnych perspektyw i stanowisk filozoficznych, zachęcając do dalszych refleksji i dyskusji nad naturą myślenia, poznania i prawdy. Natomiast intencjonalność kolektywna generuje wiele interesujących, ważnych i aktualnych pytań badawczych w filozofii społecznej i w ramach nauk społecznych. Rozważając zarówno intencjonalność jednostkową, jak i kolektywną, możemy lepiej zrozumieć naturę ludzkiego myślenia i działania, zarówno na poziomie indywidualnym, jak i społecznym. W ten sposób badania nad intencjonalnością otwierają przed nami szeroki wachlarz możliwości poznawczych i badawczych, inspirując nas do dalszych badań i refleksji nad naturą ludzkiego umysłu i społeczeństwa.

BIBLIOGRAFIA

- Brentano, F. (1999). *Psychologia z empirycznego punktu widzenia*. Przeł. W. Galewicz. Warszawa: PWN.
- Davidson, D. (1982). Rational Animals. *Dialectica*, 36(4), 318–327.
- Harman, G. (1999). Intentionality. W: W. Bechtel, & G. Graham (Red.), *A Companion to Cognitive Science* (s. 602–610). Oxford: Blackwell Publishing Ltd.
- Husserl, E. (1901). *Logische Untersuchungen*. Zweiter Teil: *Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*. *Logical Investigations*, Vol. 2. Halle: Max Niemeyer Verlag.
- Husserl, E. (1975). *Idee czystej fenomenologii i fenomenologicznej filozofii*. Księga I. Przeł. D. Gierulanka. Warszawa: PWN.
- Meinong, A. (1904). Über Gegenstandstheorie. W: A. Meinong (Red.), *Untersuchungen zur Gegenstandstheorie und Psychologie* (s. 1–51). Leipzig.
- Ingarden, R. (1931). *Das literarische Kunstwerk*. Bine Untersuchung aus dem Gremgebiet der Ontologie. *Logik und Literaturwissenschaft*. Halle: Niemeyer.
- Ingarden, R. (1987). *Spór o istnienie świata*. T. I–III, wyd. 3. Warszawa: PWN.
- Kriegel, U. (Red.). (2013). *Phenomenal Intentionality*. Oxford: Oxford University Press.
- Kripke, S. (1988). *Nazywanie a konieczność*. Warszawa: Wydawnictwo PAX.
- Pautz, A. (2013). Does Phenomenology Ground Mental Content? W: U. Kriegel (Red.), *Phenomenal Intentionality* (s. 194–234). Oxford: Oxford University Press.
- Perler, D. (Red.). (2001). *Ancient and Medieval Theories of Intentionality*. Leiden: Brill.
- Perler, D. (2002). *Theorien der Intentionalität im Mittelalter*. Frankfurt am Main: Vittorio Klostermann.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Searle, J. (1990). Collective Intentions and Actions. W: P. Cohen, J. Morgan, & M.E. Pollack (Red.), *Intentions in Communication*. Cambridge, Mass.: Bradford Books, MIT Press.
- Sellars, W. (1956). Empiricism and the Philosophy of Mind. W: H. Feigl, & M. Scriven (Red.), *Minnesota Studies in the Philosophy of Science*. Vol. I (s. 253–329). Minneapolis, MN: University of Minnesota Press.
- Tomasello, M. (2014). *Natural History of Human Thinking*. Cambridge – Massachusetts – London: Harvard University Press.

- Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. Oxford: Oxford University Press.
- Tuomela, R., & Miller, K. (1988). We-Intentions. *Philosophical Studies*, 53, 367–389.
- Twardowski, K. (1965). O treści i przedmiocie przedstawień. W: K. Twardowski, *Wybrane pisma filozoficzne*. Przeł. I. Dąmbska (s. 3–91). Warszawa: PWN.

Paweł Fortuna

Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0002-0633-4453>

Zbigniew Wróblewski

Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0003-4477-6903>

<https://doi.org/10.35765/slowniki.422>

Dusza

Streszczenie

DEFINICJA POJĘCIA: Dusza (gr. ψυχή [*psyché*], łac. *anima*) – pojęcie, które w różnych systemach filozoficznych, religijnych i kulturowych jest odnoszone do zasady życia, zasady jednoczącej i porządkującej funkcjonowanie roślin, zwierząt, ludzi lub całego świata.

ANALIZA HISTORYCZNA POJĘCIA: Można wskazać trzy typy pojmowania duszy: (a) potoczne (D1) – powiązane z tradycją ludową i wiedzą potoczną; (b) religijne (D2) – oparte na tradycjach mityczno-religijnych; (c) filozoficzne (D3) – kategoria poznawcza będąca przedmiotem naukowej refleksji.

UJĘCIE PROBLEMOWE POJĘCIA: Od początku filozofii nowożytnej obserwujemy stopniowe eliminowanie pojęcia „dusza” z debaty naukowej, teologii oraz języka potocznego na rzecz takich pojęć jak „psychika”, „umysł”, „jaźń”, „tożsamość” czy „osobowość”. Wykluczeniu D3 sprzeciwili się przedstawiciele psychologii neoscholastycznej, niektórzy badacze sfery nieświadomości, a także psycholodzy, dla których religijność była centralnym punktem refleksji.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Aktualnie najmocniejszy status można przypisać D1, a naj słabszy D3, jako pojęciu uznawanemu przez świat nauki za ontologicznie puste i epistemologicznie zbędne. Refleksja dotycząca duszy ożywa wraz z debatą nad unikalnością człowieka prowadzoną w kontekście transhumanizmu, posthumanizmu oraz antropomorfizacji systemów opartych na sztucznej

inteligencji. Znaczenie D1 motywuje do badań antropologii potocznej, opracowania typologii przekonań powiązanych z D1 oraz poznania uwarunkowań procesu ustalania znaczeń przypisywanych D1.

Słowa kluczowe: dusza, monizm, dualizm, trializm, antropologia potoczna

Definicja pojęcia

Dusza (gr. ψυχή [*psyché*], łac. *anima*) – pojęcie, które w różnych systemach filozoficznych, religijnych i kulturowych jest odnoszone do zasady życia, zasady jednoczącej i porządkującej funkcjonowanie różnych typów bytów, takich jak rośliny, zwierzęta, ludzie lub cały świat (dusza świata). W zależności od bytu duszy przypisuje się funkcje (władze): wegetatywne, zmysłowe, rozumne. Wieloznaczność pojęcia wynika z bogatej tradycji kulturowej i religijno-teologiczno-filozoficznej, co zaznacza się na poziomie języka. Duszę często utożsamia się z duchem (gr. πνεῦμα [*pnéuma*] lub νοῦς [*noús*]; łac. *spiritus* lub *animus*) jako niematerialną i nieśmiertelną substancję (esencję, samodzielnie istniejący byt) mogącą trwać po śmierci ciała fizycznego i mającą swój los w zaświatach lub innym stanie bytu. Na płaszczyźnie antropologicznej pojmuje się ją jako symbol ukrytego w człowieku wymiaru istnienia, głębi emocjonalnej i duchowej, wykraczania poza doświadczenie zmysłowe oraz związku z istotą wyższą. W tym ujęciu pojęcie duszy bywa powiązane ze świadomością i samoświadomością, tożsamością, wartościami moralnymi, sumieniem i wolną wolą. Jest obecne w tradycji ludowej i psychologii potocznej, systemach religijnych oraz w debacie naukowej, choć współcześnie dusza w takim ujęciu to kategoria krytykowana i kontestowana. W antropologii chrześcijańskiej pojęcie duszy jest kształtowane na fundamencie Objawienia (Stary i Nowy Testament) oraz poddane teologiczno-filozoficznej interpretacji w aspekcie jej natury (duchowa substancja), genezy (kreacjonizm) i funkcji (poznawcze, wolitywne, transcendentne).

Stosując kryterium genetyczne w porządkowaniu bogactwa znaczeń powiązanych z kategorią „dusza”, można wskazać trzy typy jej pojmowania: (a) potoczne (D1) – powiązane z tradycją ludową i psychologią potoczną ujmowanie duszy w perspektywie pierwszoosobowej jako niematerialnego czynnika koniecznego do wyjaśnienia funkcjonowania człowieka i zjawisk przyrody; (b) religijne (D2) – oparte na tradycjach mityczno-religijnych przypisywanie duszy znaczenia transcendentnego wraz z refleksją dotyczącą jej genezy, kontaktu człowieka z duchami, sferą nadprzyrodzoną, istotą wyższą oraz istnienia po śmierci biologicznej; (c) filozoficzne (D3) – pojmowana w perspektywie trzecioosobowej kategoria poznawcza będąca przedmiotem refleksji naukowej

i filozoficznej. Wymienione wersje znaczeniowe pojęcia „dusza” można również uszeregować według nasycenia elementami teoretycznymi – od teorii naiwnej (D1; psychologia potoczna), przez pojęcia częściowo uteoretyzowane (D2; np. w ramach dyskursu religijnego bez teologii wykorzystującej schematy pojęciowe filozofii), aż do teorii w sensie ścisłym (D3; teorie filozoficzne, teologiczne) – oraz kolejności ich nabywania w rozwoju osobniczym, czemu odpowiada uszeregowanie od wczesnodziecięcego, naiwnego pojmowania duszy (D1), aż do zreflektowanego ujęcia osób dorosłych (D2, D3).

Analiza historyczna pojęcia

Kulturowe źródła pojmowania duszy

W wierzeniach plemiennych podstawowy zbiór własności przypisywanych potocznemu pojęciu duszy (D1) zawiera następujące elementy: (a) nieokreślona siła życiowa, która występuje we wszystkich istotach żywych, a nawet w całej przyrodzie (np. Aborygeni); (b) ścisły związek (np. Jakuci) lub odrębność duszy i ciała (np. Huronowie); (c) niewidzialne stany psychiczne (np. pragnienia), które są przyczyną widzialnych zachowań (np. plemiona Hmong); (d) istnienie duszy bez ciała (np. Santalowie) (Frazer, 1925/1969). W kulturach pierwotnych można wyróżnić trzy źródła idei D1: (1) doświadczenie życia zawarte symbolicznie w tchnieniu (dusza – oddech) – proces oddychania tłumaczy życie biologiczne i jest traktowany jako znak wewnętrznej egzystencji, która może być rozwijana poza wymiarem fizycznym (np. w kulturze sumeryjskiej, mezopotamskiej, egipskiej, greckiej). Ścisłe powiązanie oddechu z duszą i duchem jest obecne na poziomie języka, np. gr. ψυχή [*psyché*] ‘dusza’ wywodzi się od czasownika ψύχω [*psychō*] ‘oddychać’ lub ‘dawać oddech’, łac. *animus* ‘dusza’ (z którego wywodzi się fr. *âme*, wł. *anima*, hiszp. *alma*) i sanskrycki *ātmán* ‘duchowy ja’ mają źródła w indoeuropejskim rdzeniu, który odnosi się do oddechu, podobnie jak ogólnosłowiańskie wyrazy *duša* i *duchъ* łączą się z wyrazem *dých* ‘oddech’; (2) nieuchronność śmierci – skończoność egzystencji ciała i trudność wyobrażenia sobie nieistnienia skłaniają do przekroczenia tych ograniczeń w idei duszy, która transcenduje normalny bieg

życia; (3) doświadczenie snu i marzeń sennych – dusza w trakcie snu ciała (jego pozornej śmierci) nadal żyje, funkcjonuje, marzy, myśli, spotyka się z bóstwami, duchami przodków; (4) pogańskie doświadczenie religijne jako źródło intuicji na temat rzeczywistości pozafizycznej, choć dyskusyjny jest pogląd, że pierwszą formą religijności był animizm, przypisujący posiadanie duszy ożywionym i nieożywionym zjawiskom przyrody (Bremer, 2010).

Religijne źródła pojmowania duszy

Religijne rozumienie duszy (D2) narodziło się w pierwotnych systemach religijnych (animizm, politeizm) i ukonstytuowało się w religiach doktrynalnych (judaizm, chrześcijaństwo, islam). W cywilizacji egipskiej wierzone, że dusza składa się z wielu aspektów: *Ka* (K'a) – indywidualna energia życiowa przypisana każdemu człowiekowi i pozostająca z ciałem po śmierci (duch-sobowtór); *Ba* (B'a) – pierwowzór duszy, nieśmiertelnej mogącej istnieć bez ciała; *Akh* (Aḥ) – aspekt duszy związany z ideą zjednoczenia z bóstwem, *Kaibit* – dusza cienia, *Ren* (Rn) – imię człowieka. W wierzeniach orfickich duszy przypisywano nieśmiertelność i wędrówkę w postaci przechodzenia przez szereg reinkarnacji w celu oczyszczenia z grzechu pierworodnego i ziemskich więzów. Koncepcja reinkarnacji pojawiła się wcześniej w hinduizmie (następnie również w buddyzmie, dżinizmie i sikhizmie), gdzie wierzy się, że dusza (*atman*) przechodzi przez szereg wcieleń zależnych od działań (*karma*) w poprzednich życiach.

W starożytnym Izraelu podkreślano zależność duszy od ciała, a starotestamentowa antropologia akcentuje jedność psychofizyczną człowieka (MacDonald, 2019). W hebrajskim tekście Starego Testamentu, będącego fundamentalnym źródłem judaizmu, słowo נֶפֶשׁ [*nefesz*] 'dusza' wiąże się z obrazem tchnienia, oznacza żądzę życia, życie samo w sobie, a także całą osobę. *Nefesz* jako pierwiastek życiodajny pochodzi od Boga, jest czynnikiem ożywiającym ciało ludzkie oraz czynnikiem organizującym jego szeroko rozumianą osobowość. Antropologia ta wskazuje trzy wymiary człowieka. Są to: בָּשָׂר [*basar*] – ciało przemijające i kruche, רוּחַ [*ruah*] – duch, który jest życiodajną siłą oraz נְשָׁמָה [*neszamah*] – oddech życia, związany z darem życiodajnego tchnienia od Boga i istnieniem życia.

Przeciwstawienie duszy i ciała pojawia się w spisanej w języku greckim antropologii nowotestamentowej. Używając słowa ψυχή [*psyché*] ‘dusza’, autorzy biblijni odsyłają do istoty życia obdarzonej świadomością moralną i przeznaczoną do nieśmiertelnej komunii z Bogiem oraz do pojęć związanych z wiecznością i zbawieniem duszy. Z kolei słowo σῶμα [*soma*] odnosi się do ciała ludzkiego zarówno w aspekcie fizycznym (nie trwała kondycja człowieka), jak i teologicznym (głębsze znaczenie związane z aspektami duchowymi i eucharystycznym). Temu przeciwstawieniu towarzyszy jeszcze inne, mianowicie πνεῦμα [*pnéuma*] ‘duch’ rozumiany jako boski Duch rozlewający się w osobie oraz σὰρξ [*sarks*] ‘ciało’ jako jego przeciwieństwo, ludzka niedoskonałość, negatywna zasada i sfera grzechu.

Wersety koraniczne prezentują pojęcie رُوح [*ruh*] ‘dusza’ jako istniejący przed narodzeniem i po śmierci dar od Boga, źródło świadomości i zdolności rozróżniania dobra od zła. Dusza może być doskonała poprzez modlitwę, praktyki religijne, studiowanie Koranu oraz prowadzenie moralnego życia. W islamie, podobnie jak w judaizmie i chrześcijaństwie, istnieje wiara w życie pozagrobowe, w którym dusza jest oceniana na podstawie uczynków i wiary.

Zgodnie z Katechizmem Kościoła Katolickiego (1994/2002) dusza jest stworzoną przez Boga nieśmiertelną częścią człowieka wprowadzaną do ciała w momencie poczęcia, a istotą ludzkiej natury jest jedność duszy i ciała:

Jedność ciała i duszy jest tak głęboka, że trzeba uważać duszę za ‘formę’ ciała; oznacza to, że dzięki duszy duchowej ciało utworzone z materii jest ciałem żywym i ludzkim; duch i materia w człowieku nie są dwiema połączonymi naturami, ale ich zjednoczenie tworzy jedną naturę (KKK, 1994/2002, 365).

Przez duszę człowiek posiada zdolność do rozróżniania dobra od zła, wolność moralną i odpowiedzialność za swoje czyny. Bezpośrednio po śmierci „każdy człowiek w swojej nieśmiertelnej duszy” (KKK, 1994/2002, 1020) podlega osądowi (sąd szczegółowy), w wyniku którego trafia do nieba, piekła lub czyśćca. Przed Sądem Ostatecznym, w którym „zostanie ostatecznie ujawniona prawda o relacji każdego człowieka z Bogiem” (KKK, 1994/2002; 1039), następuje zmartwychwstanie wszystkich zmarłych w ponownej jedności duszy i ciała.

Filozoficzne źródła pojmowania duszy

Na przestrzeni dziejów zarysowały się następujące główne sfery problemowe dotyczące filozoficznego ujęcia duszy (D3) (Judycki, 2010): istnienie, geneza, indywidualność, substancjalność, natura oraz związek z ciałem (problem psychofizyczny). W sporze o istnienie duszy ścierają się dwa stanowiska: (a) dusza nie istnieje i nie ma eksplanacyjnej potrzeby stosowania tego pojęcia w wyjaśnianiu natury człowieka (materializm, naturalizm) oraz (b) dusza istnieje i racjonalnie można wykazać (bezpośrednio w doświadczeniu fenomenologicznym lub pośrednio w argumentacji filozoficznej), że pojęcie to jest niezbędne do wyjaśnienia funkcjonowania człowieka. Dyskusja wokół problemu genezy duszy uwzględnia m.in. hipotezy sformułowane przez św. Augustyna: (a) traducjanizm (dusze dzieci pochodzą z duszy rodziców); (b) kreacjonizm (dusze są stwarzane bezpośrednio przez Boga); (c) preegzystencja (dusze istnieją przed połączeniem z ciałem ludzkim); (d) powody zarodkowe (łac. *rationes seminales*; dusze są stwarzane w zarodku ludzkim i aktualizowane w czasie). Zwolennicy indywidualizmu twierdzą, że dusza jest nierozzerwalnie związana z jednym organizmem i ma swoje własne cechy oraz tożsamość, natomiast w ujęciu holistycznym dusza jest traktowana jako część większej całości – duszy świata – lub jest emanacją duszy Boga. Substancjalistyczne ujęcie duszy zakłada, że istnieje ona jako odrębna od ciała i innych elementów rzeczywistości substancja (istota, esencja), natomiast niesubstancjalistyczne, że jest funkcją, cechą lub aspektem powiązanim z życiem organizmu, źródłem sił życiowych, siedliskiem pierwotnych dynamizmów, siłą porządkującą i podłożem ludzkich przeżyć. Określając naturę, D3 bierze się pod uwagę takie jej własności jak prostota, duchowość i niezniszczalność/nieśmiertelność. Debata nad prostotą duszy dotyczy jej pojedynczości, złożoności oraz podzielności. Jej duchowość jest definiowana przez jej niematerialność i służy wyjaśnianiu wyższych aktów psychicznych (np. poznania, miłości, abstrakcyjnego ujmowania relacji), natomiast niezniszczalność (nieśmiertelność) dotyczy zdolności egzystowania po śmierci ciała.

Różnice między przedstawionymi stanowiskami są wyraźne w kontekście debaty dotyczącej relacji duszy i ciała (MacDonalds, 2019). Od początków filozofii zarysowują się trzy główne stanowiska: monizm,

dualizm i hylemorfizm (niekiedy traktowany jako forma dualizmu). Monizm zakłada jedność podstawowej substancji lub zasady rzeczywistości. Można wyróżnić monizm materialistyczny oraz monizm spirytualistyczny. W wersji materialistycznej (fizykalistycznej) monizm redukuje całość bytu do zorganizowanej materii. Monizm idealistyczny uwzględnia wyłącznie aspekty niematerialne. Źródła monizmu odnajdujemy u presokratyków, np. Tales z Miletu uznawał duszę za siłę poruszającą, Leukippos i Demokryt za składnik materii, a Heraklit z Efezu utożsamiał naturę duszy ze zmiennością prazasady (ognia). Ideę tę rozwijali m.in. George Berkeley (idealizm subiektywny), Gottfried Leibniz (idealizm pluralistyczny) oraz Georg Hegel (monizm absolutny).

Antropologia dualistyczna przyjmująca separację duszy od ciała pozostaje pod wpływem filozofii Platona i Kartezjusza. Platon ukazywał człowieka jako istotę wewnętrznie podzieloną, w której to, co duchowe, przeciwstawiał temu, co cielesne. W jego antropologii człowiek to dusza rozumna mająca naturę boską i wieczną, natomiast śmiertelne ciało jest narzędziem jej działania. Samoistna, niematerialna i nieśmiertelna dusza istniała przed ciałem, a jej połączenie z nim (pojmowanym jako jej „więzienie”) jest tymczasowe i niekonieczne (śmierć jest ich rozdzieleniem). Według Platona dusza jest trójdzielna: rozumna (*noús*) – wyróżnia człowieka spośród żywej przyrody, jest niezbędna do poznania prawdy, rozumienia idei i osiągnięcia mądrości (można ją utożsamiać z umysłem); zmysłowo popędliwa (*thymos*) – posiadają ją również zwierzęta i jest odpowiedzialna za uczucia oraz zmysłowo pożądliva (wegetatywna, *epithymia*) – obejmuje pragnienia i jest obecna także w roślinach.

O ile Platon i inni starożytni oraz średniowieczni filozofowie dzielili byty na żyjące (posiadające duszę) i nieżyjące, o tyle Kartezjusz wprowadził podział na istoty świadome i nieświadome. W jego mniemaniu dusza to samodzielna bytowo substancja – rzecz myśląca (*res cogitans*), natomiast ciało to rzecz rozciąglą (*res extensa*). Punktem ich związania i współdziałania jest szyszynka. Kartezjusz jako pierwszy równał duszę z umysłem, który utożsamiał z myśleniem, a myślenie ze świadomością. Całkowicie oddzielił aktywność umysłową od struktury i funkcjonowania organizmu biologicznego, który zredukował do sfery zmysłowej. W ujęciu ontologicznym ciało człowieka jest złożoną strukturą maszynową, podobną do organizmów innych zwierząt. Kartezjusz dokonał

odwrotu od koncepcji duszy jako zasady życia, ponieważ według niego ciało nie wymaga duszy jako zasady życia, rośliny nie posiadają duszy wegetatywnej, a zwierzęta są czystymi mechanizmami. Późniejsze koncepcje D3 jako *res cogitans* rozwijano się w kierunku desubstancjalizacji duszy i jej funkcjonalizacji: dusza nie jest substancją, lecz zespołem funkcji psychicznych (całość życia psychicznego), czynności poznawczych, moralnych i estetycznych.

Hylemorfizm, zakładający istnienie duszy jako formy substancjalnej ciała, która żyje w jedności z nim, wprowadził Arystoteles, a później rozwijał św. Tomasz z Akwinu. Według Stagiryty dusza to akt pierwszy ciała organicznego mającego życie w możności. Dusza i ciało są ze sobą nierozzerwalnie związane – dusza nie jest ciałem, nie istniała przed nim i nie jest obdarzona wiecznym ruchem. Ciało bez duszy jest martwe, tak jak dusza bez organów ciała nie przejawia swych funkcji. Jako nieautonomiczna zasada organizująca materię do życia stanowi zbiór funkcji behawioralnych i umysłowych: dusza roślin jest odpowiedzialna za reprodukcję i odżywianie, dusza zwierząt umożliwia postrzeganie i poruszanie, odżywianie, wzrost i reprodukcję, natomiast dusza człowieka jest racjonalna, ma część aktywną i pasywną. Wymienione typy duszy nie oznaczają wielości dusz, lecz raczej aspekty jednej duszy człowieka, która zawiera duszę wegetatywną i zmysłową. Święty Tomasz z Akwinu również rozumie duszę jako formę substancjalną ciała, pierwszy i autonomiczny akt jego istnienia. Dusza nie tylko organizuje materię, lecz tworzy ją sobie i nadaje jej określone istnienie. Jako zasada życia cielesnego (dusza wegetatywna), zmysłowego (dusza zmysłowa) i intelektualnego (dusza rozumna) posiada władze, które się z nią nie utożsamiają (intelekt, wolę oraz pamięć) i jest zdolna do działań, które mają aspekty niematerialne (poznanie, decyzje, miłość). Według Akwinaty dusza, mimo ścisłej jedności z ciałem, jest nieśmiertelna, dlatego nie ciało jest jej źródłem, lecz zasada duchowa – Stwórca (kreacjonizm). Po śmierci biologicznej dusza kontynuuje swoje istnienie jako substancja niekompletna, która potrzebuje ciała do pełnienia wszystkich swoich funkcji w ramach jedności z nim (św. Tomasz z Akwinu podkreślał znaczenie wiary w zmartwychwstanie ciała).

Wersją hylemorfizmu, rozwijaną w głównym nurcie antropologii chrześcijańskiej, jest dualizm kompozycyjny (różny od dualizmu platońskiego i kartezjańskiego). W tym podejściu jedność duszy i ciała jest tak

ściła, że dusza traktowana jest jako forma ciała we wszystkich przejawach życia człowieka począwszy od funkcji fizjologicznych, a na aktach intelektualnych kończąc. Ze względu na najwyższe funkcje duszy stosuje się określenie „dusza rozumna”, a z uwagi na niematerialną naturę tych aktów – „dusza duchowa”.

Ujęcie problemowe pojęcia

Spór o istnienie duszy jest jedną z najbardziej fundamentalnych i długotrwałych debat w filozofii i nauce. Można przewidywać, że będzie ona trwała tak długo, jak długo ludzie będą zastanawiać się nad ontologiczną podstawą swojej natury i tożsamości. Większość badaczy w obliczu braku dowodów na istnienie duszy opowiada się za rezygnacją z tego pojęcia jako przedmiotu poznania naukowego – naturalizm (metodologiczny) jest aktualnie pozycją domyślną w naukach przyrodniczych. Trudność weryfikacji hipotezy duszy nie zmienia jednak faktu, że jest to pojęcie głęboko zakorzenione w systemach samowiedzy wielu osób na całym świecie, wpływając na ich tożsamość, życie duchowe i religijne, decyzje moralne, budowanie relacji czy twórczość. To skłania do prowadzenia systematycznych, interdyscyplinarnych badań przekonania ludzi na temat duszy, która, nawet jeśli jest rozpatrywana jedynie jako „użyteczna iluzja”, pozostaje kategorią egzystencjalnie istotną.

Spór o status duszy w filozofii i nauce

Od początków filozofii nowożytnej obserwujemy stopniowe eliminowanie pojęcia „dusza” z debaty naukowej, teologii oraz języka potocznego na rzecz takich pojęć jak „psychika”, „umysł”, „jaźń”, „tożsamość” czy „osobowość”. Wpłynęły na to: (a) trudność zdefiniowania D3 jako pojęcia wieloznacznego, obciążonego konotacjami kulturowymi, historycznymi i teologicznymi. Przekłada się to na problematyczność jej ujmowania w procesie badawczym w kategoriach zmiennej; (b) materializm metodologiczny – nauki przyrodnicze oraz wyłoniona z filozofii psychologia empiryczna skupiają się na mierzalnych aspektach rzeczywistości, postulują wolność od spekulacji metafizycznych i odrzucają hipotezę

duszy, kierując się niemożnością jej pomiaru; (c) brak empirycznych dowodów – w obliczu odkryć w neuropsychologii duszę uznaje się za kategorię zbędną, archaiczną (określa się ją jako mit, iluzję, szlachetne kłamstwo), czego konsekwencją jest przyjęcie hipotezy roboczej, że dusza nie istnieje; (d) sekularyzacja – wzrost akceptacji opartego na naturalizmie światopoglądu naukowego, odwrócenie się społeczeństwa od *sacrum*, podawanie w wątpliwość elementów nadprzyrodzonych i nieśmiertelności (Kugelmann, 2023).

Od końca XIX w. popularne jest określenie wprowadzone przez Friedricha A. Langego „psychologia bez duszy” (niem. *Psychologie ohne Seele*). Do eliminacji D3 z psychologii empirycznej w znaczącym stopniu przyczyniły się dzieła Franza Brentana i Williama Jamesa, choć symbolicznie duszę wykreślił James R. Angell na spotkaniu Amerykańskiego Towarzystwa Psychologicznego w 1910 r. Determinacja do obrony D2 osłabła z kolei w teologii chrześcijańskiej, co jest przejawem dystansowania się wobec koncepcji filozoficznych odbiegających od współczesnych przekonań ontologicznych (por. Possenti, 2007).

Wykluczeniu D3 sprzeciwili się przedstawiciele psychologii neoscholastycznej, niektórzy badacze sfery nieświadomości, a także psycholodzy, dla których religijność była centralnym punktem refleksji. Psychologię neoscholastyczną rozwijano od drugiej połowy XIX do połowy XX w. w celu integracji odkryć nauk przyrodniczych z myślą tomistyczną. Kierunek nadał jej kard. Désiré-Joseph Mercier, dla którego doktryna o duszy była kwestią rozumu, nie wiary. Oprócz niego D3 w XX w. uznawali m.in. Hubert Gruender SJ, Thomas Verner Moore czy Stephan Strasser. Carl G. Jung stosował pojęcie „dusza” na oznaczenie sposobu, w jaki jednostka odnosi się do swoich wewnętrznych procesów psychicznych (wewnętrzna osobowość). Interesujący wkład w myślenie o niej wniósł James Hillman (1996), który zaproponował, by duszę pojmować metaforycznie, nie jako substancję, lecz symbol i punkt widzenia będący refleksyjnym momentem między człowiekiem i zdarzeniami. W odróżnieniu od reprezentantów psychologii naukowej niedefiniowalność duszy traktował nie jako słabość, lecz siłę na wzór pojęć podstawowych w fizyce, biologii czy filozofii, które stanowią fundament ludzkiego myślenia o świecie (np. materia, życie, przyroda czy energia).

Anihilacji D3 w debacie naukowej towarzyszy przeformułowywanie pytań filozoficznych (antropologicznych) na historyczno-kulturowe,

socjologiczne, a następnie na biologiczne i informatyczne. Sprzeciwiają się temu niektórzy filozofowie i psychologowie. W krytyce podkreślane jest zagrożenie redukcjonizmem, naturalizmem i mechaniczmem, a więc tendencją do rozwiązywania problemów związanych z umysłem na polu fizyki, biologii lub chemii (Possenti, 2006). Podkreśla się również, że mimo iż teza materialistyczna jest poparta wynikami badań empirycznych, to jednak żadna z tradycyjnych teorii wyjaśniających relację duszy i ciała (z idealizmem i panpsychizmem włącznie) nie zniknęła z debaty dotyczącej problemu psychofizycznego (Nannini, 2004). W związku z tym, że materializm nie jest jednolitym programem badawczym, a konsensus naukowców w zakresie odrzucenia hipotezy duszy jest utrzymywany bardziej przez zaprzeczenie jej istnieniu niż zbiór pozytywnych rezultatów, uważa się, że przekonanie o nieistnieniu duszy można rozumieć nie jako odkrycie, lecz założenie świata nauki, i postuluje się, by hipotezę duszy traktować tak, jak inne hipotezy poddawane debacie, udoskonalaniu i testowaniu (Baker & Goetz, 2011). Wątpliwości budzi również stosowanie narzędzi nauk szczegółowych do badania D3. Poznanie tego typu fenomenów wymyka się podejściu empirycznemu, a z faktu, iż coś nie jest postrzegane w obrębie badań empirycznych, nie wynika, że to coś nie istnieje (Possenti, 2007). W rozstrzyganiu problemów antropologicznych postulowany jest pluralizm metodologiczny łączący podejście empiryczne i spekulatywne, zwłaszcza że wśród współczesnych naukowców rozważających hipotezę duszy znajdziemy zarówno biologów (Ettore Boncinelli), psychologów (Robert Kugelmann), lingwistów (Mark C. Baker), i filozofów (Richard Swinburne czy Alvin Plantinga).

Mimo podjętej przez Karla Poppera i Johna C. Ecclesa nieudanej próby wprowadzenia D3 (rozumianej w duchu kartezjańskiego dualizmu interakcyjnego) do badań neurobiologicznych, na gruncie nauk kognitywnych formułowana jest obecnie hipoteza, w myśl której koncept ten (inspirowany tradycją arystotelesowską) może być pojęciem pomocniczym w badaniach mózgu (Bremer, 2010). Chodzi o wykorzystanie „syntetyzujących” treści tradycyjnego pojęcia D3 (dusza rozumna i zmysłowa), a więc łączenie różnych stanów mentalnych i biologicznych w jedną zasadę. Ich jedność (filozoficznie wyjaśniana pojęciem duszy) znajduje empiryczne potwierdzenie, w związku z czym można ostrożnie przypisać D3 funkcję heurystyczną w badaniach, jednocześnie uznając,

że niektóre z przypisywanych jej własności (nieśmiertelność, niematerialność) są poza zasięgiem nauk empirycznych.

Krytyczna ocena D3, związana z przyjętą w nauce perspektywą trzecioosobową, natrafia na fenomenologiczne doświadczenie ludzi (perspektywa pierwszej osoby) – postulat „nauka bez duszy” nie jest tożsamy z postulatem „świat bez duszy”. Z badań przeprowadzonych w różnych kontekstach kulturowych wynika, że w istnienie duszy wierzy między 50% a 90% osób dorosłych (np. Halman et al., 2008; Richert & Harris, 2008). Wiadomo również, że duszę od umysłu spontanicznie oddzielają zarówno dzieci, jak i dorośli. Zdaniem Kugelmana (2023) nauka nie może ignorować znaczenia D1 w codziennym doświadczeniu, praktykach kulturowych, języku, a także pojmowaniu siebie, innych ludzi i przyrody. Jak podkreśla, dusza jest kategorią wciąż zakorzenioną w sposobie rozumienia świata i przejawia się w języku jako rzeczownik (dusza), przymiotnik (duchowy) oraz czasownik (ang. *doing the soul* ‘stawać się duszą’; ang. *ensoul, souling* ‘napęłniać duszą’, ‘ożywiać’).

Dusza-rzeczownik odnosi się do centralnej lub najgłębszej części istoty ludzkiej, duchowej esencji żywego organizmu – z badań wynika, że esencjalizm jest fundamentalną i uniwersalną cechą ludzkiej kognicji (Bloom, 2004). Dusza-przymiotnik wskazuje na głęboki wymiar ludzkiego istnienia, tożsamości, miłości i pasji. Odnosi się do wyższych czynności umysłowych, wykraczających poza percepcję zmysłową i reakcje emocjonalne, związanych z rozumieniem etyki, moralności czy sztuki (np. jest obecna w kontekście muzyki soul, której przesłaniem jest poszukiwanie transcendencji, pocieszenia, uzdrowienia i duchowego wymiaru bycia częścią czegoś większego, a także wyrażanie wiary, nadziei i miłości). Z kolei dusza-czasownik odnosi się do samoaktywności, podtrzymującego tożsamość sposobu bycia w świecie i stawania się sobą poprzez włączanie do indywidualności tego co inne (np. według Maurice’a Merleau-Ponty’ego dusza jest związana z żywym, dynamicznym procesem bycia w świecie, aktualizuje się w działaniu, ożywiając i nadając osobie formę, a poprzez definiowanie znaczenia sytuacji przyczynia się do kreacji sensu).

Mimo sekularyzacji współczesny człowiek poszukuje prawdy na temat duszy. Zwracał na to uwagę Jung (1933), przewidując, że zaprzeczanie jej tradycyjnemu pojmowaniu grozi poczuciem pustki wewnętrznej, alienacji i dezorientacji. Jego zdaniem może to prowadzić do rozłączenia

człowieka z jego prawdziwą naturą i utraty autentycznego związku z głębszymi wartościami, a brak zrozumienia potrzeb duchowych grozi osłabieniem poczucia sensu życia. Znaczenie samorozumienia i integralności tożsamości dla zdrowia psychicznego i poczucia spełnienia jest podkreślane w wielu koncepcjach psychologicznych. W ich kontekście podtrzymywanie koncepcji duszy jest pielęgnowaniem wewnętrznego źródła jedności i spójności myśli, uczuć i działań oraz ochroną przed rozproszeniem tożsamości i poczuciem wewnętrznego chaosu. Symptodem zapełniania pustki po duszy jest aktualnie obserwowany rozwój duchowości niereligijnej. Przekonanie o istnieniu odrębnej od umysłu, nieśmiertelnej D1 (*in form God-centered or cosmic-spiritual beliefs*) nie tylko ułatwia rozpatrywanie swojego życia w kontekście transcendencji, ale również wiąże się z uznaniem niezależności duszy od sił fizycznych, biologicznych, psychicznych i społecznych. Przeświadczenie o „nietykalności” D1 daje poczucie nienaruszalnej, wewnętrznej wolności, co pozwala ludziom chronić wewnętrzną niezależność od tortur, głodu i bólu, indoktrynacji czy stygmatyzacji. Właśnie z tego powodu ludzie odwołują się do duszy jako bazowej, antropologicznej kategorii w obliczu naruszania godności podczas wojen, migracji, konfrontacji z biedą i chorobami czy problemem uzależnień. Dodatkowo antropologiczną refleksję przywołującą pojęcie duszy prowokują toczone w wielu krajach debaty dotyczące kary śmierci, aborcji i stosowania procedur *in vitro*, a także konfrontacja z antropomorfizowanymi systemami opartymi na sztucznej inteligencji.

Przekonania na temat duszy jako przedmiot badań

Obecność duszy jako kategorii istotnej w życiu ludzi skłania przedstawicieli wielu dyscyplin do popartej badaniami empirycznymi refleksji nad współczesnym pojmowaniem D1. Nadrzędnym przedmiotem analiz są struktury pojęciowe zorganizowane wokół wiedzy jednostki na temat ontologicznej podstawy własnej natury (w kulturze Zachodniej mogą to być dusza, duch, ciało i umysł), które można traktować jako przejaw antropologii potocznej (AP; Fortuna et al., 2023a). Przekonania człowieka na temat D1 i jej powiązań z innymi aspektami definiującymi jego naturę i tożsamość są wykorzystywane w codziennym życiu jako potoczna rama pojęciowa i eksplanacyjna.

O powszechności i różnicowaniu przekonań powiązanych z AP przekonują wyniki badań nad kształtowaniem przez dzieci i dorosłych koncepcji tego, kim są, przekonaniami jednostek dotyczącymi roli duszy, mózgu i umysłu w realizacji ich funkcji biologicznych, psychobiologicznych i psychicznych, lokalizacją duszy, umysłu i jaźni w ciele, a także psychologiczną podstawą decyzji moralnych (Bloom, 2004; Richert & Harris, 2008). Systematyczna refleksja nad nimi obejmuje: (1) źródła AP, (2) różnice indywidualne w zakresie pojmowania D1 i innych aspektów ontologicznej podstawy własnej natury oraz (3) funkcje tego typu zdroworozsądkowych mniemań w życiu jednostki.

AP wyłania się jako rezultat myślenia w kategoriach esencjalizmu psychologicznego pozostającego pod wpływem indywidualnego, fenomenologicznego doświadczenia i transmitowanych kulturowo treści na temat ontologicznej podstawy natury człowieka. Dowiedziono, że we wczesnym dzieciństwie formuje się intuicyjny dualizm, na co wskazuje pojmowanie przez dzieci stanów i bytów psychicznych jako ontologicznie odmiennych od obiektów fizycznych i rzeczywistych wydarzeń (Bloom, 2004). Niezależnie od kontekstu kulturowego dzieci rozumieją, że ciała i umysły działają według różnych zasad przyczynowych, zakładając, że jeden – i tylko jeden – umysł odpowiada za zachowania jednego – i tylko jednego – ciała. Przypuszcza się, że wczesnodziecięce wrażenie dualizmu opiera się na zdolności wyobrażania sobie tego, co jest nieobecne lub nierealne, przeżywaniu snów i refleksji nad obrazami mentalnymi, a także wnioskowaniu o stanach mentalnych własnych i innych ludzi.

Problem duszy wcześniej czy później pojawia się w rozwoju osobniczym jako przedmiot rozważań, ponieważ – jak pisze Czaja (2006) – „«pytanie o duszę», o jej istnienie bądź nieistnienie, o takie czy inne jej rozumienie, jest dla każdej antropologii (tej naukowej i tej przyjmowanej na codzienny użytek) pytaniem podstawowym” (s. 26). Odbiór różnicowanych poglądów prowokuje mniej lub bardziej świadome i w różnym stopniu podlegające rygorowi krytycznej analizy rozważania natury filozoficznej i psychologicznej nadające kształt przekonaniom powiązanym z AP. Znaczącą rolę w tym procesie odgrywa treść narracji wyjaśniających niematerialny aspekt świata; eksploracja tożsamości wraz z rozważaniami filozoficznymi ukierunkowanymi na zrozumienie swojego wnętrza; wiara religijna wraz ze skupieniem się na kwestiach życia

po śmierci i kontaktu z wyższą istotą; rozumowanie teleologiczne sprzyjające odnajdywaniu celu i sensu życia; doznanie miłości i cierpienia (sytuacje graniczne), a także doświadczenia duchowe i transcendentne będące elementem praktyk medytacyjnych, kontemplacji, doznań mistycznych czy kontaktu ze sztuką.

Z badań wyłaniają się trzy typy AP: (a) antropologia jednodzielna, którą ujawniają naiwni moniści opierający ontyczną naturę własnego człowieczeństwa na cielesności (monizm fizykalistyczny) lub pierwiastku umysłowo-duchowym (monizm spirytualistyczny; idealizm); (2) antropologia dwudzielna, typowa dla naiwnych dualistów przekonanych o istnieniu ciała oraz umysłu/duszy jako odrębnych esencji oraz (3) antropologia trójdzielna, deklarowana przez naiwnych trialistów odróżniających ciało, umysł oraz duszę i na ich podstawie wyjaśniających własne funkcjonowanie. Naiwny trializm wyłania się poprzez odróżnienie duszy od umysłu. W badaniach uchwycono to różnicowanie na poziomie ontologicznym oraz funkcjonalnym. Ujawniono, że dzieci w USA potrafią dostrzegać specyficzną funkcję duszy jako duchową, a nie poznawczą czy biologiczną, natomiast wierzący w istnienie duszy dorośli z USA i Polski na ogół wyrażali przekonanie, że dusza pojawia się wcześniej niż umysł, jest bardziej stabilnym konstruktem w ciągu życia oraz istnieje dłużej od niego po biologicznej śmierci (Richert & Harris, 2008). Badani wiązali z umysłem takie funkcje jak: rozwiązywanie problemów, myślenie, pamięć, odróżnianie dobra od zła, natomiast z duszą siłę życia, życie pozagrobowe, łączność z wyższą istotą i duchowość.

W ramach podobnej AP ludzie mogą udzielać różnych odpowiedzi na pytania „czym jest dusza?” i „po co jest dusza?”. D1 jest opisywana dosłownie (niematerialna esencja, siła życiowa) i/lub poprzez metafory nawiązujące do zjawisk przyrody (ogień, tchnienie), techniki (ster, akumulator), niewidzialności (przezroczysty człowiek), fizyki (energia, pole magnetyczne), wnętrza (autonomiczne wnętrze), procesów psychicznych (świadomość), a także określeń zaczerpniętych z nauki Kościoła (iskra boża, ziarno nieśmiertelności) (Czaja, 2006). Dostrzeżono, że naiwni moniści mogą traktować umysł i duszę jako funkcje ciała (mózgu), natomiast naiwni dualiści i trialiści pojmują umysł jako funkcję duszy lub uznają ją za konstrukt „nabudowany” na umyśle. Ponadto osoby wierzące w istnienie duszy można uszeregować na „kontinuum duchowości duszy” w zależności od siły przekonania, że D1 istnieje

przed narodzinami, przetrwa śmierć biologiczną, jest źródłem siły życiowej i natury duchowej (ang. *spiritual essence*) człowieka, a także przyczynia się do zdolności osoby do życia po śmierci i nawiązania kontaktu z istotą wyższą (Richert & Harris, 2008).

Znaczenie indywidualnego pojmowania D1 przejawia się w obszarach: (a) określania przynależności do kategorii „człowiek” poprzez wskazanie duszy jako składnika ontologicznej struktury człowieka („Dusza to «ja», bo «ja» nie jestem ciałem – kiedy umrę, to ciało jest, ale mnie tam nie ma”; Czaja, 2006; s. 241); (b) kształtowania podmiotowości i tożsamości, gdy D1 jest traktowana jako kluczowy czynnik definiujący jaźń i nadający osobiste znaczenie doświadczeniom; (c) wyjaśniania funkcjonowania człowieka, w tym źródeł chorób – naiwni trialiści, w porównaniu z monistami i dualistami, przypisują duszy większą rolę w procesach biologicznych (życie, wzrost), psychobiologicznych (np. odczuwanie głodu, potrzeby snu) i psychologicznych (percepcja, pragnienia, emocje i przekonania); (d) wnioskowania moralnego – identyfikacja koncepcji D1 badanych pozwala przewidzieć ich decyzje etyczne niezależnie od deklarowanej wiary, ponieważ przekonanie o istnieniu duszy wpływa na rozważania dotyczące moralności czynów, np. na etyczną ocenę decyzji dotyczących aborcji, komórek macierzystych, czy eutanazji; osoby wierzące w istnienie umysłu i duszy uzależniają od nich nadawanie statusu moralnego innym jednostkom (np. humanoidnemu robotowi; Fortuna et al., 2023b); (e) religijności – lepsze różnicowanie duszy i umysłu współwystępuje z silniejszą religijnością i przekonaniem, że dusza jest osobną od umysłu „instancją”, która suwerennie „zarządza” ocenami moralnymi oraz transcendencją; wiara w istnienie duszy wpływa na treść rozważań dotyczących nieśmiertelności i poczucia kontaktu ze zmarłymi (Kugelmann, 2023); (f) duchowości – chrześcijanie oceniają własne uduchowienie wyżej niż buddyści wykazujący podobnie silną wiarę w istnienie duszy, którzy jednak częściej niż chrześcijanie przypisywali duszy funkcje poznawcze; (g) budowania relacji – wielu ludzi odwołuje się do D1, wyjaśniając doświadczenie głębokiej, emocjonalnej relacji z innymi osobami (np. poprzez postępowanie się takimi sformułowaniami jak „bratnia dusza”, „spotkać drugą połowę swojej duszy”); (h) twórczości – z D1 łączony jest zarówno proces tworzenia, jak i odbioru sztuki (np. „Every artist dips his brush in his own soul, and paints his own nature into his pictures” – Henry Ward Beecher).

Refleksja systematyczna z wnioskami i rekomendacjami

Aktualnie najmocniejszy status można przypisać D1 jako kategorii obecnej w codziennej refleksji „człowieka z ulicy”, a najslabszy D3 – jako pojęciu uznawanemu przez znaczącą część świata nauki za ontologicznie puste i epistemologicznie zbędne. Refleksja dotycząca duszy ożywa wraz z debatą nad unikalnością człowieka prowadzoną w kontekście transhumanizmu, posthumanizmu oraz antropomorfizacji systemów opartych na sztucznej inteligencji. Problem tożsamościowy (kim jestem?) i antropologiczny (kim jest człowiek?) aktywizuje pytanie odnoszące się do AP: co sprawia, że jestem człowiekiem? Zarówno postulowane przez zwolenników transhumanizmu biotechnologiczne udoskonalanie człowieka (projektowanie postczłowieka), jak i posthumanistyczna idea upodmiotowienia wszelkich form życia (postantropocentryczna wspólnota życia) kolidują z antropocentryczną wizją człowieka jako nadrzędnego bytu wobec innych istnień. Z badań wynika, że osoby o przekonaniach antropocentrycznych uzależniają przypisywanie statusu moralnego istotom nie-ludzkim i postaciom poddanych cyborgizacji od nadawania im umysłu i duszy, przy czym skłonność ta słabnie wraz ze wzrostem siły mniemań (Fortuna et al., 2023b).

Zdaniem Kugelmanna (2023), postulującego przywrócenie koncepcji duszy do dyskursu akademickiego, istotne jest nie tylko to, jak człowiek „tworzy duszę”, ale jak na bazie założeń o jej istnieniu „how meaning is made, how lives become, how we suffer, love, live, and die” (s. 276). Znaczenie, jakie ludzie przypisują D1, motywuje do jej eksploracji w obrębie AP. Priorytetowym zadaniem jest identyfikacja czynników psychokulturowych wpływających na kształt AP. Z badań, których uczestnikami były dzieci, wynika, że doświadczenie dualizmu pojawia się niezależnie od kulturowej transmisji idei religijnych, ale nie wiadomo, czy jest to zjawisko uniwersalne. Poza tym stosowane w językach narodowych pojęcia „dusza”, „duch” i „umysł” nie są kategoriami uniwersalnymi, a ich rozumienie zależy od konkretnej etnopsychologii i etnofilozofii. Przyjmowanie etnopsychologicznych kategorii Zachodu jako conceptualnej podstawy badań naukowych może prowadzić do błędu perspektywy (np. w kulturze chińskiej podmiotem ontologicznym nie jest osoba, lecz wielopokoleniowa rodzina), dlatego realizacja badań w innych kulturach

wymaga więc opracowania metod, które nie narzucają opartych na tradycji chrześcijańskiej przekonań dotyczących ontologicznej podstawy własnej natury.

Potwierdzona heterogeniczność grupy osób wierzących w istnienie D1 rodzi konieczność opracowania typologii AP wraz z poznaniem psychologicznych uwarunkowań procesu ustalania znaczeń przypisywanych D1 oraz czynników wpływających na kształt architektury AP. Interesująca jest analiza powiązania AP z tradycją filozoficzną i religijną, jak i prądami światopoglądowymi i narracjami popkulturowymi. Można przypuszczać, że biorąc pod uwagę przyjmowaną relację hierarchiczną między esencjami wśród naiwnych trialistów, ujawnią się naiwni spirytualiści (nadrzędność duszy), naiwni mentaliści (nadrzędność umysłu) i naiwni somatyści (nadrzędność ciała). Istotnym obszarem przyszłych badań jest także określenie funkcji, jakie D1 czy, szerzej, AP pełni w życiu człowieka w kontekście wiary i poza nim, np. wpływ na procesy atrybucji (np. Jak umarli mogą przekazywać wiadomości żywym?), teorie umysłu (np. Jak to jest być martwym?), nabywanie pojęć (np. konceptualizowanie nadprzyrodzonych właściwości zmarłych), kształtowanie przekonań (np. na temat etiologii chorób i angażowania się w zachowania prozdrowotne), rozumowanie teleologiczne (np. postrzeganie swojego życia jako zaprojektowanego dla celu, który musi zostać osiągnięty przed śmiercią), formułowania ocen moralnych (np. decyzji medycznych w zakresie diagnozy i procesu leczenia), a także doświadczanie podziwu dla piękna czy wspólnotowości.

BIBLIOGRAFIA

- Baker, M.C., & Goetz, S. (2011). Introduction. W: M.C. Baker & S. Goetz (Red.), *The soul hypothesis. Investigations into existence of the soul* (s. 1–20). New York: Continuum International Publishing Group.
- Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.
- Bremer, J. (2010). Pojęcie duszy w naukach kognitywnych. *Filozofia chrześcijańska*, 7, 37–63.
- Czaja, D. (2006). *Dusza. Figury wyobraźni i gry językowe*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

- Fortuna, P., Wróblewski, Z., & Gut, A. (2023a). Dusza w potocznej antropologii: przyczynek do identyfikacji typów ontologicznej architektury naiwnego spirytualizmu. W: T. Maziarka (Red.), *W poszukiwaniu duchowości człowieka* (s. 111–132). Kraków: Wydawnictwo CCPres.
- Fortuna, P., Wróblewski, Z., Gut, A., & Dutkowska, A. (2023b). The relationship between anthropocentric beliefs and the moral status of a chimpanzee, humanoid robot, and cyborg person: the mediating role of the assignment of mind and soul. *Current Psychology*, 43. DOI: 10.1007/s12144-023-05313-6.
- Frazer, J.G. (1969). *Złota gałąź*. Przeł. H. Krzeczkowski. Warszawa: Państwowy Instytut Wydawniczy.
- Hillman, J. (1996). *Samobójstwo a przemiana psychiczna*. Przeł. D. Rogalski. Warszawa: Wydawnictwo KR.
- Judycki, S. (2010). Istnienie i natura duszy ludzkiej. W: S. Janeczek (Red.), *Antropologia* (s. 121–177). Lublin: Wydawnictwo KUL.
- Jung, C.G. (1933). *Modern man in search of a soul*. Przeł. W.S. Dell & C.F. Baynes. New York: Harcourt Brace Jovanovich.
- Katechizm Kościoła Katolickiego* (1994/2002). Poznań: Wydawnictwo Pallottinum.
- Kugelmann, R. (2023). *Soul in soulless psychology*. Cambridge: Cambridge University Press.
- McDonald, P.S. (2003/2019). *Psyché. Dzieje pojęcia*. Przeł. I. Kania. Biblioteka Kwartanika „Kronos”. Warszawa: Fundacja Augusta hr. Cieszkowskiego.
- Possenti, V. (2007). Spór o jedność człowieka i wyzwania nowego naturalizmu. Dusza – umysł – ciało. W: A. Maryniarczyk & K. Stepien (Red.), *Dusza, umysł, ciało. Spór o jedność bytową człowieka* (s. 97–145). Lublin: Polskie Towarzystwo Tomasza z Akwinu.
- Richert, R.A., & Harris, P.L. (2008). Dualism revisited: body vs. mind vs. soul. *Journal of Cognition and Culture*, 8(1–2), 99–115. DOI: 10.1163/156770908X289224ht.

Ewa Odoj

Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0003-0821-9317>

<https://doi.org/10.35765/slowniki.484>

Problem oddziaływania psychofizycznego

Streszczenie

DEFINICJA POJĘCIA: Problem oddziaływania psychofizycznego dotyczy pytania, jak możliwa jest interakcja pomiędzy stanami mentalnymi a fizycznymi. Rozwiązanie zależy od tego, jak pojmie się naturę umysłu.

ANALIZA HISTORYCZNA POJĘCIA: Najważniejsze historyczne stanowiska w sprawie interakcji umysł-ciało to dualizm substancjalny oraz redukcyjny fizykalizm (behawioryzm logiczny, eliminatywizm, teoria tożsamości typów).

UJĘCIE PROBLEMOWE POJĘCIA: Niepowodzenie projektów o charakterze fizykalizmu redukcyjnego wynika przede wszystkim z pomijania fenomenalnego aspektu przeżyć świadomych (*qualia*). Z tego powodu najpopularniejsze dziś stanowiska dotyczące natury stanów mentalnych mają charakter nieredukcyjnego fizykalizmu. Wikłają się one jednak w zarzut epifenomenalizmu, zgodnie z którym stany mentalne, choć nie są sprowadzalne do stanów fizycznych, to jednak nie mogą oddziaływać przyczynowo w świecie fizycznym. Innym ważnym stanowiskiem jest funkcjonalizm, jednak mierzy się on z podobnymi zarzutami, co stanowiska konkurencyjne.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Współczesna refleksja nad problemem interakcji umysł-ciało powinna wziąć pod uwagę wyniki badań kognitywnych, m.in. na temat natury i genezy świadomości fenomenalnej.

Słowa kluczowe: problem psychofizyczny, koncepcje umysłu,
redukcyjny fizykalizm, nieredukcyjny fizykalizm,
funkcjonalizm

Określenie obszaru badań

Problem oddziaływania psychofizycznego, czyli interakcji pomiędzy umysłem a ciałem, jest centralnym zagadnieniem filozofii umysłu. Rozróżnienie na umysł i ciało ma swoje pierwotne źródło w doświadczeniu siebie jako świadomej, autonomicznej jaźni („ja”), która wydaje się niesprowadzalna do materialnego ciała. Jaźń jawi się jako podmiot stanów mentalnych, do których przynależą przykładowo takie zjawiska jak spostrzeganie, myślenie, chcenie, czucie, wyobrażanie. Pojawia się pytanie, czy i jak niematerialny i działający w sposób autonomiczny i wolny umysł może wpływać i podlegać wpływom materialnego ciała, które pozostaje pod ścisłym panowaniem praw przyrody. W tym duchu historycznie po raz pierwszy problem psychofizyczny wyraźnie postawiony został przez Kartezjusza, jako problem oddziaływania pomiędzy substancją duchową (łac. *res cogitans*) a cielesną (łac. *res extensa*), a następnie stał się jednym z centralnych zagadnień filozofii XVII wieku (m.in. dla Malebranche’a, Leibniza, Spinozy). Rozwój nauk przyrodniczych przyczynił się do podkreślenia innego aspektu problemu psychofizycznego: pytania o syntezę naszego potocznego wyobrażenia i mówienia o zjawiskach mentalnych, zwanego w filozofii umysłu psychologią potoczną (ang. *folk psychology*), z obrazem nas samych, jaki wyłania się z nauk przyrodniczych, w którym ujęci jesteśmy w sposób naturalistyczny, tzn. wyłącznie jako elementy świata przyrody. Z tej perspektywy w XX i XXI wieku, przede wszystkim w analitycznej, anglosaskiej filozofii umysłu, problem psychofizyczny nazwany został problemem umysł-ciało (ang. *mind-body problem*) lub nawet umysł-mózg (ang. *mind-brain problem*) w celu podkreślenia roli centralnego układu nerwowego w omawianym problemie.

Filozofowie umysłu starają się określić pewien zespół cech (predykatów) charakterystycznych dla stanów mentalnych i odróżniających je od stanów fizycznych (materialnych). W tradycji dualistycznej, której najważniejszym przedstawicielem jest Kartezjusz, podkreślano radykalną różnicę ontyczną umysłu i ciała, a za elementarną cechą stanów mentalnych uznawano ich nieprzestrzenność (w odróżnieniu od przestrzennej materii). Dziś wskazuje się na prywatność, intencjonalność i świadomość stanów mentalnych, choć żadna z tych własności nie jest niekwestionowalna (Bayne, 2022, s. 6–23). Prywatność (subiektywność)

stanów mentalnych oznacza, że są one dostępne poznawczo w bezpośredni, uprzywilejowany sposób wyłącznie dla „ja” doświadczającego owego stanu. Dla wszystkich innych podmiotów stany te są dostępne pośrednio (np. poprzez obserwację wyrazu twarzy osoby doświadczającej bólu lub przez opis). Takiej asymetrii nie ma w przypadku dostępu poznawczego do stanów fizycznych, ponieważ są one potencjalnie dane różnym osobom w taki sam sposób. Za Franzem Brentanem przyjmuje się, że charakterystyczną cechą stanów mentalnych, odróżniającą je od stanów fizycznych, jest ich intencjonalność, czyli skierowanie na coś (odniesienie do czegoś). Akty myślenia i sądzenia, pragnienia, emocje, spostrzeżenia itp. odnoszą się do czegoś innego od nich samych, do jakiegoś przedmiotu (niekoniecznie istniejącego realnie), zdarzenia lub stanu rzeczy. Kolejną cechą zdarzeń mentalnych jest ich świadomy charakter. Jest to z jednej strony intuicyjnie najbardziej oczywista ich własność, z drugiej – najtrudniejsza do opisu. Istnieją różne rodzaje świadomości (np. świadomość minimalna lub samoświadomość), jednak z punktu widzenia problemu psychofizycznego najważniejszy jej typ określa się jako świadomość fenomenalną. Nie sposób zdefiniować ją w sposób jednoznaczny i niecyrkularny, w związku z czym w literaturze z zakresu filozofii umysłu wskazuje się na jej intuicyjne znaczenie, używając frazy „jak to jest” (ang. *what it's like*) (Nagel, 1997). Podmioty świadome doświadczają pewnej specyficznej jakości przeżyciowej (określanej też jako *qualia*, w liczbie mnogiej *qualia*), towarzyszącej ich własnym stanom świadomym, dzięki której wiedzą, jak to jest np. postrzegać kolor czerwony, jak to jest doświadczać swędzenia itp. Wiedza ta jest nieosiągalna inaczej niż poprzez własną świadomość fenomenalną (por. Hutto, 2011).

Nauki przyrodnicze – przynajmniej te bazujące na fizyce klasycznej – dostarczają nam obrazu świata materialnego jako podlegającego deterministycznym prawom przyrody (por. Penrose, 1997). Określone przyczyny z konieczności wywołują określone skutki, a każde zdarzenie w świecie materialnym ma swoje przyczynowe wyjaśnienie w innym zdarzeniu o charakterze materialnym. Tymczasem przednaukowe, ale za to elementarne intuicje psychologii potocznej podpowiadają, że mamy do czynienia z czterema rodzajami oddziaływań przyczynowych: mentalno-fizycznych, fizyczno-mentalnych, mentalno-mentalnych oraz fizyczno-fizycznych. W problemie psychofizycznym kluczową rolę odgrywa

pierwszy z nich. Doświadczamy tego, że świadome „ja” może w sposób niezdeteterminowany, poprzez akty wolnej woli oddziaływać przyczynowo na przynależne mu ciało (np. poprzez podjęcie decyzji o podniesieniu ręki). Inaczej mówiąc, doświadczenie potoczne podpowiada nam, że akty mentalne mogą w sposób niezdeteterminowany wywoływać pojawienie się nowego łańcucha przyczyn w świecie fizycznym. Pojawia się pytanie, w jaki sposób pogodzić ze sobą te dwie, jak się wydaje niezgodne ze sobą, perspektywy – naukową i potoczną. Problem ten wyrazić można za pomocą trylematu, czyli zbioru takich trzech twierdzeń, z których każde wydaje się prawdziwe, a równocześnie koniunkcja jakiegokolwiek pary z nich pociąga konieczność odrzucenia trzeciego twierdzenia:

1. Zjawiska mentalne różnią się od zjawisk fizycznych, tzn. nie dają się w pełni wyjaśnić w terminach zjawisk fizycznych.
2. Zjawiska mentalne oddziałują przyczynowo na zjawiska fizyczne.
3. Świat fizyczny jest przyczynowo zamknięty, tzn. każde zdarzenie w świecie fizycznym ma swoją konieczną przyczynę w innym zdarzeniu w świecie fizycznym.

Teza (3) stanowi elementarne założenie nauk przyrodniczych, a (1) i (2) zdają relację z naszego doświadczenia potocznego. (1), (2) i (3) są niekompatybilne, ponieważ: (1) i (2) implikują, że świat fizyczny nie jest przyczynowo zamknięty; (1) i (3) implikują, że co prawda zjawiska mentalne istnieją, ale nie oddziałują przyczynowo na zjawiska fizyczne; z kolei (2) i (3) implikują, że nie istnieją zjawiska mentalne różne od zjawisk fizycznych, ponieważ pierwsze redukują się do drugich. Rozwiązanie problemu psychofizycznego poprzez zanegowanie (1) i (2) charakteryzuje stanowiska materialistyczne (zwane też naturalistycznymi lub fizykalistycznymi), a odrzucenie tezy (3) proponują stanowiska dualizmu interakcyjnego (np. Kartezjusz), których specyfiką jest podkreślanie tezy (2). Z kolei za odrzuceniem (2) opowiadają się zarówno przedstawiciele dualizmu paralelnego, przyjmujący równoległość przebiegu zdarzeń mentalnych i zdarzeń fizycznych, przy równoczesnej negacji możliwości interakcji między nimi (np. Malebranche), jak i przedstawiciele współczesnych stanowisk materializmu epifenomenalnego, którzy twierdzą, że co prawda zjawiska mentalne istnieją, ale nie oddziałują na zjawiska fizyczne, choć same podlegają wpływom zjawisk fizycznych (Bremer, 2010, s. 30–32).

Analiza historyczna pojęcia

Historycznie dominującym poglądem, charakterystycznym dla epoki nowożytnej, był dualizm substancji. Zgodnie z typowym dla tego okresu założeniem, iż porządek poznania wyznacza porządek ontyczny, myśliciele nowożytni, począwszy od Kartezjusza, przyjmowali, że muszą istnieć dwa rodzaje radykalnie różniących się od siebie substancji – ciała i umysły, ponieważ w poznaniu dane są nam dwa radykalnie różne rodzaje własności – fizyczne (przede wszystkim przestrzenność) i umysłowe (przede wszystkim nieprzestrzenność). Dla dominacji stanowiska dualizmu substancji w dziejach filozofii istotne były również inspiracje religijne, ponieważ współgra ono z religijną koncepcją duszy. Za klasycznego dualistę substancji przyjmuje się Kartezjusza. Twierdził on, że człowiek jest swoistym aglomeratem ciała (materii), które ma charakter mechaniczny, oraz nieśmiertelnej i zdolnej do istnienia bez ciała duszy (umysłu).

Problem przyczynowości psychofizycznej na gruncie dualizmu substancji wydaje się szczególnie trudny do rozwiązania (albo wręcz nierozwiązywalny), ze względu na podkreślanie w nich radykalnej różnicy ontycznej umysłu i ciała oraz przypisywanie tym substancjom przeciwnych własności. Stanowiska dualistyczne, których przedstawiciele przyjmują, że choć interakcja ta jest niezwykle trudna do wyjaśnienia (albo wręcz ma charakter niewyjaśnialnej tajemnicy), to jednak jest ona faktem (zarówno w wymiarze oddziaływania mentalno-fizycznego, jak i fizyczno-mentalnego), zwane są dualizmem interakcyjnym. Próba przezwyciężenia problemu oddziaływania psychofizycznego, przy równoczesnym zachowaniu tezy o ontycznej odrębności substancji duchowej i substancji cielesnej, jest propozycja dualizmu paralelnego. Zwolennicy tego stanowiska przyjmują albo że istnieją tylko dwa rodzaje skutecznych oddziaływań – mentalno-mentalne i fizyczno-fizyczne, albo że w ogóle nie dochodzi do jakichkolwiek bezpośrednich oddziaływań. Stanowisko to popularne było w czasach nowożytnych wśród kontynuatorów filozofii Kartezjusza. Przedstawiciele dualizmu paralelnego tego okresu zakładali, że za korelację zmian w porządku fizycznym i porządku umysłowym odpowiada Bóg, który albo te zmiany każdorazowo wywołuje (np. Malebranche), albo u początku istnienia świata przewidział wszelkie zmiany, jakie się w nim wydarzą, i ustanowił ich

harmonię (np. Leibniza). Tym samym nowożytny dualizm paralelny płacił bardzo wysoką cenę za swoją propozycję uporania się z problemem interakcji – przyjmował mianowicie bardzo mocne założenie metafizyczne, dotyczące istnienia i natury Boga, które domaga się filozoficznego uzasadnienia (Iwanicki, 2012, s. 37–41, 48–51; por. Kim, 2011, s. 31–60).

Dualizm substancji neguje fundamentalne założenie nauk przyrodniczych i z tego powodu w XX wieku został wyparty przez stanowiska fizykalizmu redukcyjnego. Niemniej jednak dualizm substancji wciąż może jawić się jako atrakcyjna koncepcja filozoficzna, ponieważ pod pewnymi względami posiada przewagę nad koncepcjami materialistycznymi o charakterze monizmu substancjalnego: lepiej radzi sobie z wyjaśnieniem doniosłego problemu tożsamości osobowej oraz pozwala na zachowanie naszych przednaukowych intuicji, dotyczących posiadania przez nas wolnej woli, dzięki której jesteśmy w stanie rozpocząć nowe ciągi przyczynowe w świecie, wbrew determinizmowi świata przyrody. Warto zauważyć, że nie jest to kwestia błaha, ponieważ przekonanie o posiadaniu przez nas wolnej woli stanowi fundamentalne założenie życia społecznego (w tym ustroju demokratycznego), w którym przypisujemy podmiotom prawa i obowiązki.

Wśród popularnych w XX wieku stanowisk o charakterze fizykalizmu redukcyjnego wymienić należy: (1) behawioryzm logiczny, (2) teorię identyczności typów, (3) materializm eliminacyjny (eliminatywizm). Zakłada się w nich, że możliwe jest takie sprowadzenie (redukcja) stanów mentalnych do stanów fizycznych, po dokonaniu którego okaże się, że problem interakcji umysł-ciało nie jest rzeczywistym problemem, gdyż oddziaływanie, do którego się odnosi, jest po prostu oddziaływaniem tego, co fizyczne, na to, co również fizyczne. Wspólne tym stanowiskom jest więc założenie metafizycznego monizmu w kwestii ostatecznej natury rzeczywistości – istnieje tylko jeden rodzaj substancji, tzn. substancje materialne.

W latach 20. i 30. XX wieku przedstawiciele behawioryzmu logicznego (np. Gilbert Ryle) – najmocniejszej postaci fizykalizmu redukcyjnego – pod wpływem pozytywizmu logicznego Koła Wiedeńskiego twierdzili, że zjawiska umysłowe można opisać w intersubiektywnie dostępnych kategoriach behawioralnych jako dyspozycje do określonych zachowań. Postulowali oni eliminację terminu „świadomość” i wszelkich

terminów o charakterze mentalnym do terminów odnoszących się do zdarzeń obserwowalnych (np. ktoś zażywa tabletkę i mówi, że źle się czuje w przypadku odczuwania bólu). Redukcja ta miała mieć charakter konieczności analitycznej (na mocy definicji) (Heil, 2020, s. 50–70). Przedstawiciele materializmu eliminacyjnego (np. Paul Churchland), wychodząc z założenia, że tylko nauka, a ostatecznie fizyka, wyznacza to, co istnieje (tzw. realizm naukowy), głosili, że język mentalistyczny należy zastąpić językiem fizykalnym, ponieważ ten pierwszy odnosi się do czegoś, co w rzeczywistości nie istnieje. Ich zdaniem psychologia potoczna to przednaukowa prototeoria, która z czasem zostanie całkowicie wyeliminowana przez naukowy opis neurofizjologii człowieka, podobnie jak teoria spalania, jako utleniania, zastąpiła XVII-wieczną hipotezę flogistonu, który uważany był za materiał ulatniający się podczas spalania. Według przedstawicieli materializmu eliminacyjnego tak, jak hipoteza flogistonu stała się całkowicie zbędna w świetle nowej teorii fizykalnej, tak samo język mentalistyczny zostanie wyparty z naszej mowy wraz z rozwojem neuronauk (Heil, 2020, s. 166–177). Kolejną, ulepszoną wersją materializmu redukcyjnego była teoria identyczności typów (ang. *type physicalism*; *type-type identity theory*). Jej przedstawiciele (np. John Smart), zainspirowani przykładami udanych identyfikacji w przyrodoznawstwie (np. błyskawica to wyładowanie elektryczne), postulowali redukcję zdarzeń mentalnych do zdarzeń fizycznych poprzez ustalenie identyczności (tożsamości) określonych typów stanów mentalnych i określonych typów stanów neurofizykalnych, np. u wszystkich ludzi ból K jest tożsamy ze wzrostem aktywności włókien C₄. Redukcja ta miała mieć charakter konieczności empirycznej (na mocy prawa przyrody). Zdaniem Smarta „doznania” (stany mentalne) i „procesy mózgowie” są dwiema różnymi nazwami, które odnoszą się do tego samego przedmiotu (analogicznie do terminów „błyskawica” i „wyładowanie elektryczne”), choć jedno to wyrażenie potoczne, a drugie – wyrażenie naukowe (Heil, 2020, s. 71–87).

Programy o charakterze redukcyjnego fizykalizmu spotkały się z ostrą krytyką. Przykładowo przedstawicielom behawioryzmu logicznego zarzucano, że postulowana przez nich redukcja jest niewykonalna, ponieważ nie każde zdarzenie mentalne łączy się z jakimś zachowaniem. Możemy wyobrazić sobie istoty całkowicie odporne na ból, które pomimo doświadczenia go nie reagują jakimkolwiek zachowaniem.

Podobnie moglibyśmy mieć do czynienia z osobą, która perfekcyjnie symuluje zachowanie kogoś, kto doświadcza bólu, choć w rzeczywistości go nie doświadcza. Eliminatywizm skrytykować można jako stanowisko samoznoszące: jeżeli eliminatywista jest przekonany, że nie istnieją przekonania, to tym samym podważa swój własny pogląd. Hilary Putnam (Putnam, 1975) przedstawił słynną krytykę teorii identyczności typów, zwaną zarzutem wielorakiej realizowalności (ang. *multiple realizability*). Teoria ta utożsamia określony typ zdarzeń mentalnych z dokładnie jednym typem zdarzeń fizycznych. Tymczasem jest wysoce wątpliwe, iż istnieje dokładnie jeden typ procesu neuronalnego, który związany jest np. ze wszystkimi przypadkami bólu u wszystkich istot odczuwających ból. Przykładowo tak prymitywne zwierzę, jakim jest ośmiornica, wydaje się również odczuwać ból, choć w przypadku tych istot realizuje się on na bazie innych procesów neuronalnych niż te, z jakimi mamy do czynienia w przypadku odczuwania bólu przez ludzi (Bayne, s. 38–55, por. Kim, 2011, s. 61–127).

Ujęcie problemowe pojęcia

Formułowane w XX wieku programy redukcyjnego fizykalizmu spotkały się z poważną krytyką, w wyniku której straciły na popularności. Najważniejsze zarzuty wobec stanowisk o tym charakterze odnoszą się do wspomnianego wyżej zagadnienia jakości stanów przeżyciowych (*qualiów*). Szczególne znaczenie w dyskusjach na ten temat miały opublikowane w drugiej połowie XX wieku artykuły Thomasa Nagela i Franka Jacksona. Ten ostatni w tekście *Czego Maria nie wiedziała* (Jackson, 2003) przedstawił słynny eksperyment myślowy, opisujący neuronaukownicę Marię zamkniętą od urodzenia w czarno-białym pokoju. Posiada ona całą dostępną wiedzę na temat fizykalnych aspektów widzenia barw, ale nigdy nie widziała kolorów innych niż czarny i biały. Po wyjściu z pokoju Maria po raz pierwszy doświadczyła widzenia barw, a więc poznała fenomenalne jakości barw (*qualia*). Intuicyjnie wydaje się, że Maria zdobyła nową wiedzę, której do tej pory nie posiadała (i nie mogła posiadać), mianowicie dowiedziała się, jak to jest spostrzegać kolory. Rozumowanie Jacksona można zrekonstruować w następujący sposób:

1. Wiedza fizykalna nie jest całą wiedzą.
2. Jeśli fizykalizm jest prawdziwy, to wiedza fizykalna jest całą wiedzą.
3. Fizykalizm jest fałszywy (Iwanicki, 2012, s. 66).

Eksperyment myślowy dot. Marii ma na celu wykazać prawdziwość przesłanki 1.

W podobnym duchu argumentuje Nagel w artykule *Jak to jest być nietoperzem* (Nagel, 1997). Wskazuje w nim, że nikt z nas nie wie, jak to jest orientować się w przestrzeni dzięki zdolności do echolokacji, jak czynią to nietoperze. Fizjolog może wiedzieć wszystko na temat nietoperzy, ale nigdy nie posiada pierwszoosobowej, doświadczalnej wiedzy na temat tego, jak to jest być nietoperzem. Zdaniem Nagela fizyczny opis świata i wszelka wiedza naukowa są zawsze niewyczerpujące, bo pomijają fenomenalny aspekt przeżyć świadomych, możliwy do poznania tylko z perspektywy pierwszoosobowej (Chalmers, 2003; Bayne, 2022, s. 133–153; Iwanicki, 2012, s. 54–63, 65–68; por. Bremer, 2005, s. 193–212).

Argumenty te zwane są argumentami z wiedzy, ponieważ broni się w nich nieredukowalnego charakteru wiedzy dot. własnych przeżyć świadomych. Równocześnie traktowane są jako mocne racje na rzecz dualizmu własności, czyli poglądu zakładającego realność i wyjątkowość zarówno własności fizycznych, jak i mentalnych. Tym samym stanowią one ważne kontrargumenty względem stanowisk o charakterze redukcyjnego fizykalizmu, którego przedstawiciele postulują redukcję własności mentalnych do własności fizycznych. W związku m.in. ze wskazaną tu krytyką projektów redukcjonistycznych popularność zyskały stanowiska klasyfikowane jako fizykalizm nieredukcyjny. Ich przedstawiciele twierdzą, że choć na sposób substancjalny istnieją jedynie ciała materialne (a zatem zakładają monizm substancjalny o charakterze materializmu), to jednak niektóre z nich posiadają obok własności fizycznych również własności mentalne, których nie da się zredukować do tych pierwszych. Tak więc na gruncie fizykalizmu nieredukcyjnego z jednej strony nie porzuca się fundamentalnego założenia przyrodoznawstwa o ostatecznej materialnej budowie świata przyrody, z drugiej natomiast nie odmawia się realności stanom mentalnym, choć równocześnie nie przypisuje się im ontycznego statusu substancji.

Nieredukcyjny fizykalizm przyjmuje trzy kluczowe tezy: 1. „Istnieją własności mentalne, które są różne od jakichkolwiek własności fizycznych”.

2. „Własności mentalne zależą od (ang. *depend on*) własności fizycznych”. 3. „Własności mentalne mają przyczynowy wpływ na wydarzenia [w świecie – E.O.]” (Baker, 2009, s. 110–111).

Jeśli chodzi o tezę pierwszą, to oznacza ona, iż nieredukcyjni fizykaliści przyjmują wspomniany wyżej dualizm własności. Ich zdaniem bytom materialnym mogą, choć nie muszą (jak jest w przypadku bytów nieożywionych), przysługiwać równocześnie dwa rodzaje własności – fizyczne i mentalne. Teza (2) jest sformułowana bardzo ogólnie, a jej właściwe dookreślenie stanowi najważniejsze i najtrudniejsze zadanie przedstawicieli różnych stanowisk o charakterze nieredukcyjnego fizykalizmu. Za Donaldem Davidsonem powszechnie przyjęło się określać wspomnianą w tej tezie relację mianem superweniencji (łac. *super* – na, ponad, *venire* – przybywać). Nazwa ta sugeruje nadbudowanie własności (lub zjawisk) drugiego rzędu, w tym przypadku mentalnych, na własnościach (zjawiskach) pierwszego rzędu, mających charakter podstawy – w tym przypadku fizycznych. Zwolennicy stosowania tego pojęcia starają się zdefiniować je w taki sposób, aby oznaczało ono relację na tyle silną, by implikowała zależność stanów mentalnych od stanów fizycznych, a równocześnie na tyle słabą, aby nie pociągała redukcji jednych do drugich. Przykładowo według Davidsona zbiór własności A superweniuje na innym zbiorze B wtedy i tylko wtedy, gdy jeśli jakiś obiekt zmienił się pod względem własności mentalnych, to zmienił się również pod względem własności fizycznych (por. Davidson, 1992, s. 175).

Teza (3) oznacza odrzucenie przez nieredukcyjnych fizykalistów epifenomenalizmu, czyli poglądu, zgodnie z którym co prawda własności mentalne istnieją i różnią się od własności fizycznych, ale są jedynie korelatami (epifenomenami) zjawisk o charakterze fizycznym (procesów neuronalnych) i nigdy nie odgrywają roli przyczynowej w świecie fizycznym. Nieredukcyjni fizykaliści stawiają sobie tym samym za cel obronę założenia obecnego w doświadczeniu potocznym i w naukowej psychologii na temat posiadania przez stany (i zdarzenia) mentalne mocy sprawczej. Ponadto większość nieredukcyjnych fizykalistów chce pozostać w zgodzie z naukowym obrazem świata (a właściwie dopełnić ten obraz teorią z zakresu filozofii umysłu), w związku z czym przyjmują za niekwestionowalne założenie o przyczynowym zamknięciu świata fizycznego, tzn. założenie, iż każde zdarzenie w świecie fizycznym ma swoją wystarczającą przyczynę w innym zdarzeniu w świecie

fizycznym (zob. teza (3) trylematu psychofizycznego przedstawiona powyżej) (Baker, 2009, s. 109–113; Heil, 2020, s. 178–184).

Program nieredukcyjnego fizykalizmu został poważnie skrytykowany przez Jaegwona Kima, który przedstawił między innymi tzw. argument z przedeterminowania (ang. *overdetermination*). Twierdzi w nim, że jeżeli określone zdarzenie mentalne, nazwijmy je M1, wywołuje inne zdarzenie mentalne – M2, to istnieje takie zdarzenie fizyczne F1, które jest podstawą, na której superweniuje M1, i równocześnie istnieje takie zdarzenie fizyczne F2, które jest podstawą dla M2. Jeżeli świat materialny jest przyczynowo zamknięty, co stanowi wspomniane już fundamentalne założenie nauk przyrodniczych, to F2 ma swoją wystarczającą przyczynę w F1. W związku z tym wydaje się, że M2 jest przedeterminowane, tzn. ma dwie przyczyny. Z jednej strony jest wywołane przez M1, ale równocześnie, jeśli M2 superweniuje na F2, to wystarczającą przyczyną F2, którą jest F1, jest też wystarczającą przyczyną M2. Zdaniem Kima, jeżeli zdarzenia mentalne nie są identyczne ze zdarzeniami fizycznymi, to każde zdarzenie mające przyczynę w zdarzeniu mentalnym jest przedeterminowane – ma dwie wystarczające przyczyny. Sytuację tę Kim porównuje do sytuacji, w której dwóch zabójców niezależnie od siebie postrzela śmiertelnie tę samą ofiarę w idealnie tym samym momencie. Jego zdaniem jest nieprawdopodobne, żeby tego typu sytuacja miała miejsce w każdym przypadku przyczynowania mentalnego (Baker, 2009, s. 114). Argumentacja Kima odnosi się do każdego rodzaju przyczynowania mentalnego. Powyżej przedstawiona była na przykładzie przyczynowania mentalno-mentalnego, ale z analogiczną sytuacją mamy do czynienia w odniesieniu do przyczynowania mentalno-fizycznego. Jeśli M oznacza decyzję, by podnieść rękę, P odpowiednie fizyczne zdarzenie neuronalne, które jest podstawą dla M, a E ruch ręki, to rozumowanie Kima można zrekonstruować w następujący sposób:

1. P jest przyczynowo wystarczające do wywołania E (KOMPLETNOŚĆ).
2. Jeśli P jest przyczynowo wystarczające do wywołania E, to nic, co jest inne od P, nie wywołuje E, chyba że E jest przedeterminowane (WYŁĄCZNOŚĆ).
3. M jest różne od (tzn. nie identyczne z) P (ODRĘBNOŚĆ).
4. Nie mamy do czynienia z przedeterminowaniem (NIEPRZEDETERMINOWANIE).
Więc:
5. M nie wywołuje E (Bayne, 2022, s. 181).

W świetle argumentacji Kima przyczynowanie mentalne wydaje się pozorne, ponieważ wszystkie zdarzenia mają swoje wystarczające przyczyny fizyczne. Jeśli rozumowanie Kima, w tym przyjęte przez niego przesłanki, są poprawne, to nieredukcyjny fizykalizm pociąga za sobą epifenomenalizm, niezależnie od deklaracji i chęci jego zwolenników. Odrzucenie tego wniosku byłoby możliwe, jeśli któraś z przesłanek nie byłaby wiarygodna. Teza o kompletności jest ściśle związana ze wspomnianym wyżej założeniem nauk przyrodniczych o przyczynowym zamknięciu świata przyrody. Teoretycznie możemy mieć do czynienia z sytuacją przedeterminowania (jak w przypadku śmierci z powodu dwóch równoczesnych postrzałów), ale takie wyjaśnienie każdego zdarzenia mentalnego wydaje się mało prawdopodobne. Przyjęcie tezy o wyłączności i nieprzedeterminowaniu wydaje się więc zasadne. Dla zachowania twierdzenia o przyczynowaniu psychofizycznym pozostaje więc odrzucić tezę o odrębności. To natomiast oznacza powrót do redukcyjnego fizykalizmu w postaci teorii, która zakłada identyczność stanów mentalnych i stanów fizycznych. Zgodnie z takim redukcjonizmem, tak jak pożar lasu jest wywołany przez wyładowanie elektryczne i piorun, ale nie jest to sytuacja przedeterminowania, bo są to dokładnie te same zjawiska, tak ruch mojej ręki jest wywołany przez fizyczne procesy neuronalne tożsame z moimi stanami mentalnymi. Wniosek z argumentacji Kima wydaje się więc następujący: albo zaakceptujemy identyczność stanów mentalnych i stanów fizycznych, albo uznamy, że nie są one identyczne, ale przyczynową sprawczość mają wyłącznie zdarzenia fizyczne. W obu przypadkach pozostaje mrzonką nasze przekonanie o możliwości wywoływania przez zjawiska mentalne (np. akty wolnej woli) nowych, niezdeterminowanych ciągów przyczynowych w świecie fizycznym (zob. teza (2) trylematu) (Bayne, 2022, s. 180–183; Kim, 2009, s. 38–46; Kim, 2011, s. 214–223; Baker, 2009, 113–116; Heil, 2020, s. 184–193).

Intensywny rozwój badań nad sztuczną inteligencją (ang. *Artificial Intelligence*, w skrócie AI) sugeruje, aby szczególną uwagę poświęcić stanowisku funkcjonalizmu, którego pierwsze sformułowania pochodzą z połowy XX wieku. Zgodnie z podstawową ideą funkcjonalizmu określony stan umysłowy należy utożsamiać z funkcją (lub rolą), jaką pełni on w odpowiednio zorganizowanym systemie. Funkcjonalisci postrzegają stany fizyczne i stany mentalne jako dwa rodzaje opisów – materialne

i funkcjonalne – jednego bytu substancjalnego. Stanowisko to zrodziło się dzięki inspiracjom zaczerpniętym z informatyki, a jego źródeł doszukać się można w pracach pioniera tej dziedziny nauki – matematyka Alana Turinga. Zaproponował on rozumienie umysłów w kategoriach maszyn liczących, a stanów mentalnych jako wyników procesów obliczeniowych mózgu. Ujęcie to stanowi podstawę tzw. funkcjonalizmu maszynowego, inaczej zwanego komputacjonizmem. Stosunek między mózgiem a umysłem pojmowany jest w tej postaci funkcjonalizmu analogicznie do relacji pomiędzy komputerem (*hardware*) a jego oprogramowaniem (*software*). Zgodnie z pierwotną ideą funkcjonalizmu, zaproponowaną przez Putnama, nie jest istotne, jakie jest materialne podłoże stanów mentalnych – system, wykonujący określone funkcje umysłowe (np. mnożenie), może być zarówno organem biologicznym, jakim jest mózg, jak i układem elektronicznym. Wynika to z przyjęcia przez funkcjonalistów tezy o wielorakiej realizowalności stanów mentalnych. W związku z tym koncepcja ta dopuszcza możliwość przypisywania własności umysłowych nie tylko ludziom, ale również innym istotom żywym, a nawet urządzeniom (np. komputerom). Na gruncie komputacjonizmu mózg utożsamia się z maszyną obliczeniową, realizującą określone programy, a więc dające się wyrazić liczbowo algorytmy, które przedstawiciele tego stanowiska utożsamiają ze stanami umysłowymi. Koncepcja ta dostarczyła bardzo owocnego paradygmatu dla badań kognitywnych, w tym badań nad sztuczną inteligencją. Komputacjonizm zakłada tzw. tezę mocnej AI, zgodnie z którą maszyny liczące mogą (przynajmniej potencjalnie) myśleć, tak samo jak istoty żywe, posiadające rozbudowane biologiczne mózgi (w odróżnieniu od tzw. tezy słabej AI, zgodnie z którą maszyny mogą jedynie symulować myślenie) (Kim, 2011, s. 129–167; Van Gulick, 2009, s. 128–131; Heil, 2020, s. 88–105).

W typowej krytyce funkcjonalizmu wskazuje się m.in., że wykonywanie określonych funkcji, rozumianych jako wywoływanie tych samych skutków, co stany mentalne, nie wystarcza do tego, abyśmy mieli do czynienia ze stanami świadomymi, które w myśl tej krytyki są fundamentalną cechą stanów umysłowych. Na takie wnioski wskazują dwa niezwykle wpływowe argumenty przedstawione w postaci eksperymentów myślowych – chińskiego pokoju autorstwa Johna Searle’a oraz chińskiego mózgu spopularyzowany przez Neda Blocka. Argumenty te podważają również tezę wielorakiej realizowalności stanów mentalnych.

Wskazują bowiem, że nie każda materia może dostarczyć podłoża do pojawienia się charakterystycznych dla stanów mentalnych cech świadomości i intencjonalności.

Przeciwko tezie mocnej AI, i tym samym przeciwko komputacjonizmowi, Searle sformułował słynny argument chińskiego pokoju. Przedstawił w nim sytuację zamkniętego w pokoju człowieka, który w najmniejszym stopniu nie rozumie języka chińskiego. Z zewnątrz podawane mu są pytania w języku chińskim, na które formułuje odpowiedź, bazując wyłącznie na książce zawierającej instrukcje, jakich znaków alfabetu chińskiego użyć w odpowiedzi na pojawienie się w pytaniu określonego tekstu w języku chińskim. Scenariusz ten ma symulować tzw. test Turinga, którego zdanie ma według funkcjonalistów świadczyć o faktycznym myśleniu określonego systemu (maszyny Turinga). Zdaniem Searle'a osoba zamknięta w chińskim pokoju może dostarczać perfekcyjnych odpowiedzi, a mimo to w najmniejszym stopniu nie rozumieć języka chińskiego. W tej sytuacji jedynie symuluje ona świadome stany mentalne związane z rozumieniem tego języka. Ponadto, pomimo operowania znakami chińskimi, czyli syntaksą tego języka, całkowicie poza jej zasięgiem jest jej semantyka, czyli odnoszenie się znaków językowych do rzeczywistości poza nimi. Osobie tej brakuje więc tego, co nazwane zostało wcześniej intencjonalnością charakterystyczną dla stanów mentalnych.

W eksperymencie myślowym wielkiego chińskiego mózgu wskazuje się, iż na mocy zasady wielorakiej realizowalności określony mentalny stan funkcjonalny może być zrealizowany przez całą populację Chin w ten sposób, iż każdy Chińczyk wykonuje takie samo zadanie, jakie jest charakterystyczne dla pojedynczego neuronu. Każdy Chińczyk dostaje pewne dane wejściowe, na które reaguje zgodnie z posiadanymi instrukcjami, które polecają mu, jakie dane wyjściowe powinien w danej sytuacji przekazać kolejnemu Chińczykowi. W myśl tej krytyki wykonywanie takich komend przez całą populację Chin odpowiada temu, jak funkcjoniści rozumieją stany mentalne. Tymczasem jednak nie jesteśmy w stanie wskazać, gdzie w tej złożonej z Chińczyków symulacji mózgu mamy do czynienia ze świadomymi stanami fenomenalnymi (*qualia*). Nie można przypisać ich żadnemu pojedynczemu Chińczykowi (tak samo jak nie możemy przypisać np. świadomego postrzegania zieleni pojedynczemu neuronowi), ani tym bardziej całej

populacji Chin, biorącej udział w tym eksperymencie. Wskazuje to, że funkcjonalne rozumienie stanów mentalnych jest niewystarczające, ze względu na pomijanie ich fenomenalnego charakteru (Braddon-Mitchell & Jackson 2007, s. 107–128).

Krytyka Searle'a i Blocka pozostaje do pewnego stopnia zbieżna ze wspomnianymi wyżej argumentami Jacksona i Nagela przeciwko redukcjonistycznemu materializmowi. Zarówno redukcja stanów mentalnych do stanów fizycznych, jak i redukcja do stanów funkcjonalnych (zwłaszcza rozumianych na sposób przyczynowo-skutkowy) pomija świadomościowy charakter stanów mentalnych i tym samym nie dostarcza wystarczającego opisu tego, czym one są (por. Van Gulick, 2009, s. 143–149). Ta niewystarczalność funkcjonalnej koncepcji umysłu sprawia, że zjawiska mentalne nie są na jej gruncie właściwie ujęte, a konsekwentnie uniemożliwia to udzielenie przez przedstawicieli tego stanowiska satysfakcjonującej odpowiedzi na problem interakcji psychofizycznej.

Funkcjonalizm ma jednak jeszcze inne trudności związane z problemem interakcji. Mianowicie, podobnie jak antyredukcyjny fizykalizm, funkcjonalizm wydaje się pociągać epifenomenalizm. W związku z tym, iż własności mentalne ujmowane są na gruncie tej koncepcji funkcjonalnie, a własności materialnego podłoża strukturalnie, to relacja pomiędzy tymi pierwszymi a drugimi określana jest jako realizacja (inaczej: implementacja, wykonanie). Tak jak program komputerowy realizowany jest na określonym systemie komputerowym, tak stany umysłowe realizowane są w przypadku człowieka na biologicznej materii mózgu. Własności funkcjonalne (program komputerowy, stany mentalne) stanowią tu własności drugiego rzędu, które realizują się na własnościach pierwszego rzędu – materialnych własnościach systemu (komputera, mózgu). Z tej perspektywy widać, w jaki sposób funkcjonalizm podlega analogicznej krytyce, jaką względem teorii superweniencji sformułował Kim: całość przyczynowej sprawczości przypada materialnym własnościom systemu. Realizowany stan funkcjonalny jest jedynie pochodną materialnych własności systemu, które odpowiadają za powstanie fizycznych ciągów przyczynowo-skutkowych. W związku z tym funkcjonalista powinien albo uznać, iż przyczynowa sprawczość mentalna jest jedynie pozorna, a więc uznać stanowisko epifenomenalizmu, albo relację pomiędzy własnościami mentalnymi i fizykalnymi ująć nie tylko jako realizację określonej funkcji, ale również jako jakiś rodzaj identyczności

(typów lub egzemplarzy). W obu przypadkach funkcjonałności nie uda się rozwiązać trylematu psychofizycznego, czyli pogodzić naszych trzech intuicji. Zamiast tego w pierwszym przypadku odrzuci tezę drugą trylematu, a w drugim przypadku – pierwszą (Van Gulick, 2009, s. 142–143; Kim, 2009, s. 46–48).

Wnioski

Problem oddziaływania psychofizycznego zależy od tego, jak pojmie się stany mentalne. Jeśli zredukuje się je do stanów fizycznych, to okaże się on pozorny, ponieważ oddziaływanie, którego dotyczy, z tej perspektywy będzie po prostu oddziaływaniem fizyczno-fizycznym. Takie ujęcie stanów mentalnych jest jednak niesatysfakcjonujące z wielu względów, choć najważniejszy z nich dotyczy pomijania świadomościowej charakterystyki stanów mentalnych, zwłaszcza przeżyciowych jakości fenomenalnych zwanych *qualiami*. Podobny problem napotykają próby zdefiniowania stanów mentalnych wyłącznie poprzez odniesienie do funkcji, jakie pełnią one w złożonym systemie. Takie ujęcia również wydają się być „ślepe” na świadomościowy wymiar stanów mentalnych – system może realizować funkcje mentalne bez typowej dla stanów mentalnych charakterystyki fenomenalnej i intencjonalnej. Jeśli natomiast przyjmie się, że stanów mentalnych nie da się zredukować, ale są to unikalne własności nadbudowane jako własności drugiego rzędu na własnościach fizycznych, to co prawda uda się zachować swoistość stanów mentalnych, ale pod poważnym znakiem zapytania stanie możliwość przyczynowej sprawczości tak pojętych stanów mentalnych. W tym ujęciu zarówno określenie relacji pomiędzy stanami mentalnymi a fizycznymi jako relacji superweniencji, jak i jako relacji realizacji (implementacji) określonej funkcji, nie pozostawia miejsca na realną sprawczość stanów mentalnych, ponieważ zgodnie z zasadą przyczynowego zamknięcia świata fizycznego całość oddziaływania przyczynowego przypada fizycznym własnościom pierwszego rzędu.

Pomimo tych trudności zdecydowana większość badaczy zajmujących się zagadnieniami z zakresu filozofii umysłu nie podejmuje prób obrony raczej historycznego już stanowiska dualizmu substancji. Część z nich, przykładowo Daniel C. Dennett, opowiada się za epifenomenalizmem,

a więc odmawia stanom mentalnym działania sprawczego. Traktują oni psychologię potoczną, w ramach której wydaje nam się, że takie oddziaływanie zachodzi, wyłącznie jako użyteczną fikcję, która pozwala przewidywać i tłumaczyć zachowania ludzi, co z kolei jest niezbędne do prowadzenia normalnego życia społecznego. Opis funkcjonalny stanów mentalnych (zwłaszcza jeśli podkreśla się w nim aspekt behawioralny) jest użyteczny z punktu widzenia rozwoju nauki i w związku z tym wciąż chętnie przyjmowany przez badaczy problematyki umysłu (nie tylko filozofów), jednak idea funkcjonalizmu zostaje zaadaptowana do pogłębiającej się wiedzy z zakresu nauk kognitywnych. Przykładowo zwraca się uwagę, że funkcjonalny poziom świadomości i intencjonalności nie pojawia się dzięki prostemu podziałowi systemu na poziom strukturalny i funkcjonalny. Wyrafinowane przejawy działania umysłu, takie jak wysoka inteligencja, intencjonalność, samoświadomość, pojawiają się dzięki integracji bardziej elementarnych podsystemów, składających się na mózg ludzki. Ideę funkcjonalizmu należy więc rozpatrywać na wielu zależnych od siebie poziomach, a nie oczekiwać pojawienia się tak zaawansowanych form funkcjonowania umysłów w pojedynczym „kroku”, który stanowić miałby całe przejście od poziomu fizykalnego do mentalnego (Van Gulick, 2009, s. 138–141; por. Bremer, 2005, s. 151–192). Współcześnie ogranicza się również tezę wielorakiej realizowalności stanów mentalnych – pojawienie się stanów świadomych wydaje się ściśle związane z własnościami biologicznej materii organicznej systemów neuronalnych. Co więcej, w naukach kognitywnych pod szyldem „ucieleśnionego umysłu” podkreśla się nieodzowną rolę ludzkiego ciała w tworzeniu i funkcjonowaniu wszelkich zjawisk mentalnych. Współczesne ujęcia umysłu odcinają się więc od kartezjańskiego dualizmu substancji i elementarnych założeń tego stanowiska. Problem świadomości, zwłaszcza jej zaawansowanych form, m.in. świadomości fenomenalnej (*qualiów*), wydaje się dziś stanowić podstawową zagadkę, której rozwiązanie jest niezbędne do udzielenia odpowiedzi na problem psychofizyczny (por. Bremer, 2005).

BIBLIOGRAFIA

- Baker, L.R. (2009). Non-Reductive Materialism W: A. Beckermann, B.P. McLaughlin, & S. Walter (Red.), *The Oxford Handbook of Philosophy of Mind* (s. 109–127). Oxford: Oxford University Press.
- Bayne, T. (2022). *Philosophy of Mind. An Introduction*. Abingdon: Routledge.
- Braddon-Mitchell, D., & Jackson, F. (2007). *The Philosophy of Mind and Cognition*. Oxford–Cambridge: Blackwell Publishing.
- Bremer, J. (2005). *Jak to jest być świadomym. Analityczne teorie umysłu a problem neuronalnych podstaw świadomości*. Warszawa: Wydawnictwo IFiS PAN.
- Bremer, J. (2010). *Wprowadzenie do filozofii umysłu*. Kraków: Wydawnictwo WAM.
- Chalmers, D.J. (2003). Consciousness and its Place in Nature. W: S.P. Stich, & T.A. Warfield (Red.), *The Blackwell Guide to Philosophy of Mind* (s. 102–142). Malden: Blackwell Publishing.
- Davidson, D. (1992). Zdarzenia mentalne. Przeł. T. Baszniak. W: D. Davidson. *Eseje o prawdzie, języku i umyśle*. B. Stanosz (Red.) (s. 163–193). Warszawa: PWN.
- Heil, J. (2020). *Philosophy of Mind. A Contemporary Introduction*. New York: Routledge.
- Hutto, D. (2011). *Consciousness*. W: J. Garvey (Red.), *The Bloomsbury Companion to Philosophy of Mind* (s. 35–53). London: Bloomsbury.
- Iwanicki, M. (2012). Dualizm psychofizyczny. Odmiiany, argumenty, zarzuty. W: M. Miłkowski, & R. Poczobut (Red.), *Przewodnik po filozofii umysłu* (s. 37–84). Kraków: Wydawnictwo WAM.
- Jackson, F. (2003). Czego nie wiedziała Maria? Przeł. T. Ciecierski. *Przegląd Filozoficzno-Literacki*, 4(6), 9–14.
- Kim, J. (2009). Mental Causation. W: A. Beckermann, B.P. McLaughlin, & S. Walter (Red.), *The Oxford Handbook of Philosophy of Mind* (s. 29–52). Oxford: Oxford University Press.
- Kim, J. (2011). *Philosophy of Mind*. Boulder: Westview Press.
- Nagel, T. (1997). Jak to jest być nietoperzem? Przeł. A. Romaniuk. W: T. Nagel, *Pytania ostateczne* (s. 203–220). Warszawa: Fundacja Aletheia.
- Penrose, R. (1997). *Makroświat, mikroświat i ludzki umysł*. Przeł. P. Amsterdamski. Warszawa: Prószyński i S-ka.
- Putnam, H. (1975). Philosophy and Our Mental Life. W: H. Putnam, *Mind, Language and Reality: Philosophical Papers*. T. 2 (s. 291–303). Cambridge: Cambridge University Press.

Van Gulick, R. (2009). Functionalism. W: A. Beckermann, B.P. McLaughlin, & S. Walter (Red.), *The Oxford Handbook of Philosophy of Mind* (s. 128–151). Oxford: Oxford University Press.

Marek Pokropski
Uniwersytet Warszawski
<https://orcid.org/0000-0002-7423-6951>
<https://doi.org/10.35765/slowniki.413>

Naturalizacja fenomenologii

Streszczenie

DEFINICJA POJĘCIA: Naturalizacja fenomenologii oznacza nieredukcyjną integrację fenomenologii, wywodzącej się z tradycji Husserlowskiej, z kognitywistyką. Celem tak rozumianej naturalizacji jest włączenie fenomenologicznych badań w obszar empirycznych badań nad świadomością.

ANALIZA HISTORYCZNA POJĘCIA: Historia naturalizacji fenomenologii sięga lat 90. XX wieku, kiedy to w kognitywistyce pojawiło się zainteresowanie takimi zjawiskami jak świadomość i ucieleśnienie procesów poznawczych. Zwrócono wówczas uwagę na to, że fenomenologia może wnieść dużo do wyjaśnienia tych zjawisk.

UJĘCIE PROBLEMOWE POJĘCIA: W części problemowej krytycznie omówione zostaną najważniejsze stanowiska teoretyczne w dyskusji na temat naturalizacji fenomenologii. Szczegółowo przeanalizowano dwie propozycje fenomenologii znaturalizowanej: neurofenomenologię, która łączy metody fenomenologiczne z metodami obrazowania aktywności mózgu, oraz fenomenologię fazy wstępnej, która fenomenologii przypisuje rolę w procesie projektowania eksperymentów.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: W części poświęconej refleksji systematycznej i wnioskowi wskazane zostaną braki dotychczasowej dyskusji na temat naturalizacji oraz możliwe kierunki kontynuacji. Jednym z takich kierunków jest umieszczenie problemu naturalizacji fenomenologii w kontekście dyskusji nad unifikacją eksplanacyjną kognitywistyki i pokazanie, że fenomenologia może zostać

zintegrowana z innymi modelami wyjaśnień, w szczególności z modelem mechanistycznym.

Słowa kluczowe: fenomenologia, Husserl, naturalizacja, kognitywistyka

Definicja pojęcia

Naturalizacja fenomenologii oznacza jej nieredukcyjną integrację z naukami kognitywnymi. Takie rozumienie naturalizacji przyjmują redaktorzy tomu *Naturalizing Phenomenology*, który zbiera szereg propozycji teoretycznych dotyczących tego zagadnienia (Roy et al., 1999). Należy podkreślić, że integracja ta ma zachodzić na poziomie wyjaśnień, tzn. celem naturalizacji jest zintegrowanie fenomenologii z wyjaśnieniami zgodnymi ze standardami nauk przyrodniczych, a także wykazanie ciągłości między fenomenologią a naturalistycznie zorientowaną kognitywistyką. Zdaniem zwolenników naturalizacji fenomenologii jej przeprowadzenie przyczyni się do postępu w wyjaśnianiu zjawisk umysłowych. Jak dotąd nie udało się sformułować zadowalającej teorii naturalizacyjnej, przedstawiono natomiast kilka interesujących propozycji, jak integracja fenomenologii z naukami kognitywnymi mogłaby wyglądać.

Podana definicja wymaga dwóch komentarzy. Po pierwsze, użyty w niej termin „fenomenologia” odnosi się do nurtu filozoficznego zapoczątkowanego przez Edmunda Husserla na początku XX wieku, w którym szczególną wagę przywiązywano do opisu i analizy aktów świadomości oraz przeżywanego doświadczenia. Celem badań fenomenologicznych jest odkrycie struktury i genezy aktów świadomości. Oprócz Husserla fenomenologami, których najczęściej wymienia się w kontekście naturalizacji, są Martin Heidegger, Jean-Paul Sartre i Maurice Merleau-Ponty. Wśród nich szczególne miejsce zajmuje ten ostatni ze swoją fenomenologią ciała, która w znaczny sposób przyczyniła się do powstania popularnej dziś w kognitywistyce koncepcji poznania ucieleśnionego. Należy zaznaczyć, że w nieco innym znaczeniu termin „fenomenologia” występuje w analitycznej filozofii umysłu, gdzie rozumie się przez nią subiektywny charakter doznawanych jakości wrażeniowych nazywanych przez niektórych *qualiami* (np. Chalmers, 1996/2014). Naturalizacja fenomenologii w tym drugim znaczeniu polegałaby na pokazaniu, w jaki sposób własności fenomenalne stanów umysłowych wynikają z procesów fizycznych zachodzących w mózgu. Dla filozofii fenomenologicznej wywodzącej się od Husserla jakości wrażeniowe nie stanowią centralnego tematu badań, nie oznacza to jednak, że nie może ona pomóc w rozwiązaniu tego problemu.

Drugi komentarz dotyczy rozumienia naturalizmu. Ogólnie rzecz ujmując, w obrębie naturalizmu możemy wyodrębnić dwa stanowiska: ontologiczne, które stwierdza, że wszystko, co istnieje, jest częścią przyrody, oraz epistemologiczne lub metodologiczne, zgodnie z którym za obowiązujące uznawane są metody badania i wyjaśniania stosowane w naukach przyrodniczych. W odniesieniu do zagadnienia naturalizacji fenomenologii naturalizm ontologiczny rozumiany jest nieredukcyjnie. Mówiąc inaczej, postulowana w definicji ciągłość między fenomenologią i kognitywistyką ma mieć charakter nieredukcyjny, czyli powinna uwzględniać własności postulowane przez fenomenologię i nie prowadzić do ich eliminacji lub zastąpienia przez własności badane w kognitywistyce. Jednak projekt naturalizacji fenomenologii skupia się bardziej na zagadnieniach metodologicznych, choć i tutaj redukcjonizm zastąpiony jest integracjonizmem. Oznacza to, że naturalizacja fenomenologii nie powinna prowadzić do porzucenia jej metod ani do wyjaśnienia umysłu wyłącznie za pomocą terminów charakterystycznych dla nauk przyrodniczych. Przeciwnie, chodzi o metodologiczne połączenie tych tradycji badawczych, o integrację wyników badań fenomenologicznych z wynikami badań nauk kognitywnych, w szczególności z badaniami z zakresu neuronauki.

Analiza historyczna pojęcia

Związki fenomenologii z naukami kognitywnymi zaczęły się na dobre zacieśniać w latach 90. XX wieku. Wcześniej fenomenologia, o ile odnosiła się do kognitywistyki, to raczej występowała w roli jej krytyczki, wskazując na słabości dominującej w tamtym czasie obliczeniowej teorii umysłu. Według tej teorii procesy poznawcze da się wyjaśnić w kategoriach przetwarzania reprezentacji symbolicznych analogicznie do funkcjonowania komputerów. Warto w tym kontekście wspomnieć prace Huberta Dreyfusa, który z pozycji Heideggerowskich krytykował projekt sztucznej inteligencji, wskazując na konieczność jej ucieleśnienia i usytuowania w świecie. Dreyfus był również jednym z pierwszych filozofów, którzy do dyskusji kognitywistycznych wprowadzili filozofię Husserla, przy czym odczytywał on Husserla jako prekursora obliczeniowej teorii umysłu (np. Dreyfus & Hall, 1982). Według tej interpretacji Husserlowskie pojęcie noematu odpowiadałoby pojęciu reprezentacji umysłowej

będącej wynikiem procesów obliczeniowych. W latach 90. ubiegłego wieku klasyczny paradygmat obliczeniowy zaczął słabnąć i pojawiło się zainteresowanie zagadnieniami takimi jak świadomość i ucieleśnienie podmiotów poznających, czyli tematami centralnymi dla fenomenologii.

Świadomość została wprowadzona do kognitywistyki jako „trudny problem” związany z tzw. luką eksplanacyjną, czyli problem, którego rozwiązanie leży poza możliwościami nauk kognitywnych, a jego ewentualnego rozwiązania może dostarczyć bliżej nieokreślona nauka przyszłości (np. Chalmers, 1996/2014). Problematyczny charakter mają, według Chalmersa, fenomenalne własności stanów umysłowych, czyli treści wrażeniowe, takie jak ból, doznanie koloru czerwonego czy zapach kawy, które, jego zdaniem, wymykają się przyczynowym i funkcjonalnym wyjaśnieniom kognitywistów. O ile fenomenologia Husserlowska nie skupia się na problemie jakości wrażeniowych, w terminologii Husserla określanym mianem dat hyletycznych, którym znaczenie nadaje intencjonalność, o tyle w oczywisty sposób może ona wnieść dużo do zrozumienia tej intencjonalnej struktury doświadczenia, a przez to do wyjaśnienia świadomości. W tym kontekście fenomenologię wprowadził do kognitywistyki Francisco J. Varela, twórca neurofenomenologii i jeden z pomysłodawców projektu naturalizacji, nazywając ją „remedium na trudny problem” (Varela 1996/2010). Varela był przekonany, że znaturalizowanie fenomenologii doprowadzi do rozwoju naukowej metody badania subiektywnych stanów doświadczenia.

Drugim ze wspomnianych zagadnień jest wpływ ucieleśnienia na procesy poznawcze, np. w jaki sposób ciało organizmu, jego budowa i zdolności motoryczne wpływają na percepcję, pamięć etc. Obok prac Jamesa Gibsona z zakresu psychologii ekologicznej to fenomenologia miała istotny wkład w tworzenie paradygmatu poznania ucieleśnionego. Fenomenologia – w szczególności późne prace Husserla, takie jak *Idee II* (1974), i *Fenomenologia percepcji* Merleau-Ponty’ego (2001) – pokazuje, jak nasze ciało związane jest ze świadomością i jak wpływa na procesy poznawcze oraz nasze bycie i działanie w świecie. Fenomenologiczna analiza ciała wskazuje na kluczową rolę wrażeń dotykowych oraz kinestetycznych (związanych z ruchem) w konstytucji zarówno doświadczenia własnej cielesności, jak i rzeczy materialnych. Struktura doświadczenia percepcyjnego ukazuje, że to ciało daje nam unikatową perspektywę postrzegania świata wokół nas. Z miejsca, w którym

zlokalizowane jest nasze ciało, rzutowane są relacje przestrzenne (np. blisko-daleko, lewo-prawo), z tej lokalizacji widzimy określony profil przedmiotów, ale to nasze ciało, dzięki swoim możliwościom motorycznym, umożliwia nam zmianę perspektywy i spostrzeżenie innych profili. Cieleśnie wchodzimy także w interakcję z przedmiotami, które wykorzystujemy jako narzędzia i które szybko stają się naturalnymi przedłużeniami naszych możliwości. Mówiąc ogólnie, zdaniem fenomenologów, nasza cielesność kształtuje świadomość, m.in. procesy percepcyjne, orientację przestrzenną, przeżywanie emocji. Wreszcie to poprzez ciało, które jest ekspresyjne i intencjonalnie działające w świecie, poznajemy inne podmioty, inne ucieleśnione świadomości.

Nie wolno przeoczyć, że oprócz świadomości i cielesności fenomenologia wnosi do kognitywistyki rozbudowaną analizę intencjonalności. Pojęcie to pochodzi z tradycji psychologii opisowej Franza Brentana i oznacza własność aktów świadomości polegającą na skierowaniu na jakiś przedmiot lub, w innej interpretacji, posiadanie przez stany umysłowe przedmiotowej treści reprezentacyjnej, poprzez którą odnoszą się do reprezentowanych przedmiotów. Jak wiadomo, intencjonalność była głównym tematem badań fenomenologii Husserla, który z jednej strony przypisywał jej funkcję nadawania sensu danym wrażeniowym, z drugiej strony wskazywał na jej różne odmiany, np. intencjonalność percepcyjną, intencjonalność emocji, intencjonalność sygnitywną w aktach mowy czy pamięciową w przypomnieniu. Na przykład mogę widzieć niedźwiedzia, mogę się go bać, ale też mogę opowiadać o tym, jak spotkałem niedźwiedzia w Tatrach, lub mogę to sobie przypominać. W Husserlowskiej fenomenologii rozpoznanie i opisanie odmian aktów intencjonalnych jest kluczowe dla zrozumienia świadomości i jej relacji ze światem. Z kolei fenomenologia Merleau-Ponty'ego (2001) rozszerza to pojęcie na ciało, bo to nasze ciało przejawia pierwotną intencjonalność działania i bycia w świecie. Kognitywistyka z osiągnięć fenomenologii w tym zakresie czerpie, choć – jak się wydaje – w niewystarczającym stopniu.

Husserl nie poprzestał na tzw. statycznej analizie intencjonalności, polegającej na jej typologii i opisie struktury, i wprowadził perspektywę genetyczną, która miała na celu opisanie genezy aktów świadomości. Sam zamysł dobrze obrazuje prosta konstatacja, że intencjonalne akty świadomości zachodzą w czasie i ulegają zmianie, np. zmienia się treść wrażeniowa, a także, że mają na nie wpływ czynniki nieświadome, takie

jak doświadczane cieleśnie pobudzenia afektywne. Z tego względu dla Husserla (1989) tak ważne było badanie świadomości czasu, na którą składają się dwie główne funkcje nazwane retencją i protencją. W skrócie: retencja jest zatrzymywaniem w świadomości tego, co właśnie przeminęło w czasie, natomiast protencja jest funkcją antycypacji tego, co ma niebawem nadejść. Retencja i protencja umożliwiają rozpostarcie aktów świadomości w czasie, zachowując jednocześnie tożsamość aktu świadomości i jego przedmiotu. Jest to kolejny ważny, a na pewno najczęściej przywoływany wkład fenomenologii do kognitywistycznych badań nad świadomością. Pojęcia wywodzące się z fenomenologii świadomości czasu były stosowane np. w badaniach nad percepcją, motoryką ciała, interakcją społeczną, a także w wyjaśnianiu zaburzeń psychicznych takich jak depresja.

Husserl i krytyka naturalizmu

Warto w tym miejscu cofnąć się do pierwszej połowy XX wieku i krótko nawiązać do stanowiska, jakie wobec filozofii naturalistycznej podzielał sam Edmund Husserl. W swoich pracach otwarcie krytykował próby redukcyjnego wyjaśnienia świadomości w sposób czysto naturalistyczny, czyli za pomocą metod i terminologii nauk przyrodniczych. Według Husserla zwolennik naturalizmu postrzega każdy przedmiot badań jako po prostu część natury, która jest rządzona przez ściśle prawa dające się opisać w języku matematyki, jak ma to miejsce w matematycznym przyrodoznawstwie. Jednak takie podejście wynika z iluzji istnienia uniwersalnej matematycznej metody naukowej, którą można by stosować w dowolnej dziedzinie, również w badaniach świadomości. Dla Husserla jest to niedorzeczność, gdyż świadomość, po pierwsze, nie jest zwykłą częścią przyrody tak, jak np. kamienie i gwiazdy, a po drugie, struktury i treści doświadczenia nie da się opisać za pomocą ścisłych praw czy to przyczynowych, czy formalnych. Dlatego do wyjaśnienia świadomości niezbędna jest metoda fenomenologiczna, zdolna do uchwycenia struktury doświadczenia i wypracowania jej nietrywialnego pojęcia.

Należy jednak podkreślić, że w krytyce naturalizmu Husserl bynajmniej nie zaprzecza związkowi świadomości z jej fizycznym, organicznym podłożem. Świadomość jest ściśle związana z ciałem, co Husserl

wykazuje np. w obszernych fragmentach drugiej księgi *Idei* (1974). Husserl nie był także przeciwny badaniom empirycznym w psychologii i psychofizyce prowadzonym np. przez Wilhelma Wundta, Gustava Fechnera czy przedstawicieli psychologii postaci. Wręcz przeciwnie, dostrzegał dużą wartość takich badań. Twierdził jednak, że psychologia empiryczna nie ma odpowiedniej podbudowy teoretycznej i dlatego ujmuje świadomość w sposób powierzchowny, przez co nie będzie w stanie jej wyjaśnić. Reforma psychologii, zdaniem Husserla, wymaga jej ugruntowania w fenomenologii. W wykładach wygłoszonych w 1925 roku, zatytułowanych *Phänomenologische Psychologie* (1962), Husserl zaprezentował projekt nowej psychologii, zwanej psychologią fenomenologiczną, którą rozumiał jako dyscyplinę pośredniczącą między fenomenologią transcendentálną a psychologią empiryczną. Zadaniem psychologii fenomenologicznej byłoby badanie istoty (struktury) aktów świadomości (np. form świadomości percepcyjnej) i dostarczanie psychologii empirycznej pojęć, których nie da się wypracować w nastawieniu naturalistycznym, a które nadałyby psychologii ugruntowanie teoretyczne.

Podsumowując, stanowisko Husserla nie jest antynaturalistyczne, tylko antyredukcjonistyczne. Husserl krytykuje redukcyjny naturalizm, który chce badać świadomość wyłącznie metodami matematycznego przyrodoznawstwa i który nie uwzględnia subiektywnego wymiaru doświadczenia. Jego zdaniem konieczne jest zreformowanie badań psychologicznych i nadanie im fenomenologicznych podstaw teoretycznych. Oczywiście nie wynika z tego, że Husserl byłby orędownikiem dzisiejszego projektu naturalizacji fenomenologii. Wydaje się raczej, że byłby przeciwny wszelkiej naturalizacji, w której fenomenologia traci swój przewodni i fundacyjny charakter. W formułowanych dzisiaj propozycjach fenomenologii znaturalizowanej tak się właśnie dzieje, a fenomenologia staje się jedną z wielu dyscyplin, które mogą coś wnieść do interdyscyplinarnych nauk kognitywnych.

Ujęcie problemowe pojęcia

W dyskusji na temat możliwości naturalizacji fenomenologii można wyróżnić trzy główne stanowiska. Pierwsze z nich zdecydowanie odrzuca taką możliwość, twierdząc, że istotą fenomenologii jest jej

transcendentalizm (np. Moran, 2013), którego celem jest rozjaśnienie warunków możliwości poznania jako takiego. Takie odczytanie fenomenologii wyklucza jej naturalizację i ustawia ją wobec nauk przyrodniczych w roli dyscypliny fundującej.

Drugie stanowisko wskazuje na możliwość współpracy fenomenologii i kognitywistyki, przy czym fenomenologia musiałaby zostać w odpowiedni sposób pozbawiona transcendentalizmu (zob. np. Pokropski, 2021). To dość ogólnie sformułowane stanowisko daje szeroki wachlarz możliwości, jak taka współpraca miałaby wyglądać od strony metodologicznej. Na przykład Dan Zahavi (2010) proponuje luźno zdefiniowaną współpracę opartą na wzajemnej inspiracji, która mogłaby się przełożyć na projekty badawcze. W podobnym duchu wypowiada się Shaun Gallagher (2017), który pisze o „wzajemnym oświeceniu” fenomenologii i kognitywistyki. Jego propozycja, nazwana fenomenologią fazy wstępnej (*front-loaded phenomenology*), wskazuje na możliwość zastosowania koncepcji fenomenologicznych na wczesnym etapie projektowania eksperymentów kognitywistycznych. Stanowiska, które bardziej rygorystycznie podchodzą do zagadnienia współpracy i relacji między fenomenologią a kognitywistyką, posługują się pojęciem ograniczeń (*constraints*), które miałyby być nakładane przez fenomenologię na kognitywistykę i odwrotnie. Natura tych ograniczeń pozostaje tematem dyskusji (zob. np. Pokropski, 2021; Roy et al., 1999). Podczas gdy jedni zakładają możliwość zmapowania struktur doświadczenia na aktywność neuronalną i proponują do tego użyć formalnych relacji, jak izo- lub homomorfizm, drudzy zastanawiają się nad nowym rodzajem wzajemnej relacji, np. neurofenomenologia posługuje się, niestety dość niejasnym, pojęciem pasaży generatywnych (*generative passages*) (np. Lutz & Thompson, 2017). Jeszcze inni (Pokropski, 2021) twierdzą, że ograniczenia dostarczane przez fenomenologię powinny być rozumiane zgodnie z dominującym w neuronauce mechanistycznym modelem wyjaśniania, czyli powinny prowadzić do ograniczenia przestrzeni możliwych mechanizmów odpowiedzialnych za wyjaśniane zjawisko przez wskazanie własności tych mechanizmów.

Stanowisk na temat naturalizacji fenomenologii dopełnia trzeci pogląd, według którego nie tyle chodzi o naturalizację fenomenologii, ile o fenomenologizację natury, czyli fenomenologiczną redefinicję pojęcia przyrody, a w konsekwencji o zmianę paradygmatu nauk

przyrodniczych (np. Gallagher, 2018). Punktem wyjścia jest Husserłowska krytyka nauk przyrodniczych, które ignorują subiektywny charakter procesów poznawczych oraz błędnie rozumieją przyrodę jako coś, co jest „gotowe” do odkrycia przez „bezosobowe” i neutralne nauki. Fenomenologia, przeciwnie, wskazuje w duchu kantowskim na aktywny udział podmiotu poznającego w konstytucji poznawanej przyrody. Mówiąc inaczej, poznawana przyroda to przyroda, która się przejawia jakimś podmiotowi lub wspólnocie podmiotów. Dlatego zdaniem zwolenników fenomenologizacji natury należy wprowadzić do jej pojęcia aspekty subiektywności. Można to zrobić, wprowadzając do słownika nauk przyrodniczych takie znaczenia jak intencjonalność, ucieleśnienie świadomości, sprawstwo (*agency*) czy usytuowanie podmiotu w kontekście sytuacji praktycznej.

Patrząc na wymienione stanowiska z perspektywy szans projektu naturalizacji, należy stwierdzić, że pierwsze stanowisko zamyka dyskusję, wskazując na niemożliwość integracji z naukami przyrodniczymi transcendentalizm. Faktem jest, że fenomenologia Husserla zwróciła się w pewnym momencie w stronę filozofii transcendentalnej (zob. np. pierwsza księga *Ideii*). Jednak zawężanie Husserłowskiego projektu do wątków transcendentalnych wydaje się nieuprawnione, gdyż pomija jego dużą różnorodność tematyczną, a także zmienność na przestrzeni lat. Ponadto nawet jeśli zgodzimy się, że naturalizacja fenomenologii Husserla jest niemożliwa do przeprowadzenia, to nie wyklucza to naturalizacji innych fenomenologii, np. fenomenologii ciała Merleau-Ponty’ego. Z kolei stanowisko trzecie wydaje się zbyt mocne, gdyż postuluje daleko idącą reformę nauk przyrodniczych, która byłaby trudna do zaakceptowania przez naukowców. Inna sprawa, że Husserłowska krytyka naturalizmu, do której odwołują się zwolennicy tej koncepcji, jest dziś w dużej mierze anachroniczna. Posługuje się wizją przyrody, która jest w pełni deterministyczna i rządzona przez uniwersalne i ścisłe prawa. Natomiast obraz przyrody, który wyłania się z nauki współczesnej, uwzględnia jej niedeterministyczny charakter (np. mechanika kwantowa). Dzisiejsza nauka wskazuje także, że obawy fenomenologów odnośnie do naturalistycznej redukcji są nieaktualne, gdyż pomysł redukcyjnej unifikacji nauki na podstawie dedukcyjno-nomologicznego modelu wyjaśniania okazał się niemożliwy do spełnienia. Ponadto współczesna kognitywistyka, głównie za sprawą koncepcji poznania ucieleśnionego,

uwzględnia w dużej mierze takie zjawiska jak ucieleśnienie, sprawstwo czy usytuowanie poznania. Wydaje się więc, że z tych trzech stanowisk to stanowisko drugie wskazujące na mniej lub bardziej rygorystyczną relację wzajemnego ograniczania ma szansę na realizację i jest podzielane przez większość zwolenników naturalizacji.

Matematyzacja

Tematem, który pojawia się często w kontekście naturalizacji fenomenologii, jest możliwość zastosowania matematyki do analizy fenomenologicznych opisów struktur doświadczenia. Zdaniem niektórych jest to wręcz warunek możliwości powodzenia tego projektu (zob. np. Roy et al., 1999). Zaletą takiego podejścia jest, że matematyczne formuły wyrażałyby własności doświadczenia w sposób ścisły i dostępny dla innych badaczy. Matematyzacja mogłaby także zapewnić przekład różnych reprezentacji występujących w fenomenologii i naukach przyrodniczych na uniwersalny język liczb.

Stanowisko samego Husserla, skądinąd matematyka z wykształcenia, wobec zastosowania matematyki w fenomenologii było niejednoznaczne. Z jednej strony obwinał matematyzację nauk przyrodniczych o przyczynienie się do kryzysu racjonalności (np. *Kryzys nauk europejskich i fenomenologia transcendentálna*) oraz podkreślał różnice między podejściem matematycznym i fenomenologicznym (np. *Idee I*). Według Husserla fenomenologia nie jest dyscypliną formalną, jak matematyka, ale materialną, to znaczy zajmuje się aktami świadomości i skorelowanymi z nimi przedmiotami doświadczenia. Natury tej korelacji, jak i treści przedmiotowych nie da się wyrazić w ścisłych formułach. Wynikałoby z tego, że zastosowanie metod matematycznych do świadomości jest błędem. Z drugiej jednak strony Husserl przyznaje, że za sukcesem współczesnej nauki stoi metoda matematyczna, w szczególności zastosowanie geometrii do opisu świata materialnego oraz formalizacja praw natury. W swoich pracach często posługiwał się również zapisami quasi-matematycznymi. Na przykład w *Badaniach logicznych* używa zapisu symbolicznego, żeby analizować hierarchiczną strukturę aktów percepcji i przypomnienia, natomiast w *Wykładach z fenomenologii wewnętrznej świadomości czasu* przedstawia przeptyw świadomości

w formie diagramów. Wydaje się zatem, że pogląd Husserla na temat związku matematyki z fenomenologią nie jest do końca spójny.

W opinii Roya, Petitota, Pachouda i Vareli (Roy et al., 1999) zdystansowane stanowisko Husserla wobec zastosowania matematyki w badaniach świadomości było konsekwencją ograniczeń nauk ścisłych jego czasów. Na przykład w czasach Husserla zastosowanie teorii układów dynamicznych ograniczało się do fizyki, natomiast współcześnie teoria ta jest aplikowana w badaniach różnych dyscyplin, m.in. w ekologii, biologii, jak i w kognitywistyce. W skrócie: teoria systemów dynamicznych zajmuje się opisem i analizą zachowania układów złożonych, a dokładniej analizą zmiany tego zachowania w czasie. Narzędzia formalne tej teorii składają się z równań różniczkowych, w których zmienne reprezentują kluczowe własności analizowanych systemów, a parametry odnoszą się do warunków zewnętrznych. Zachowanie badanego systemu ujęte jest formalnie przez zbiór takich równań. Można je również przedstawić wizualnie w postaci trajektorii w n -wymiarowej przestrzeni stanów systemu, gdzie n oznacza liczbę zmiennych. Zdaniem Vareli to właśnie teoria układów dynamicznych jest drogą do naturalizacji fenomenologii, ponieważ modele dynamiczne mogą dobrze reprezentować zmieniające się w czasie własności subiektywnego doświadczenia.

Wybrane propozycje metodologiczne: neurofenomenologia

Neurofenomenologia została zaproponowana przez Francisca Varełę (1996), chilijskiego biologa i neuronaukowca. Głównym celem tego projektu było opracowanie metody badawczej łączącej fenomenologię i neuronaukę, która pomogłaby rozwiązać „trudny problem” świadomości, czyli wyjaśnić związek pomiędzy subiektywnym doświadczeniem a leżącymi u jego podstaw procesami neuronalnymi. Varela, czerpiąc inspiracje z fenomenologii Husserla (np. odwołując się do pojęcia redukcji fenomenologicznej), stwierdza, że możliwe jest rzetelne badanie struktur i własności doświadczenia z perspektywy pierwszoosobowej, a opisy tego doświadczenia można połączyć z danymi pochodzącymi z metod neuroobrazowania, np. elektroencefalografii (EEG). Co ważne, neurofenomenologia interesuje się ogólnymi strukturami doświadczenia

powtarzającymi się w przeżyciach wielu podmiotów a nie opisami indywidualnych doświadczeń i ich jakościowym charakterem. Jak stwierdza Varela:

nowatorstwo mojej propozycji polega na tym, że metodyczne ujęcia pierwszoosobowe powinny być integralnym składnikiem oceny propozycji neurobiologicznych, a nie tylko przypadkową czy heurystyczną informacją. Właśnie z tego powodu zdecydowałem się na opis tej sytuacji za pomocą hipotezy, zgodnie z którą oba ujęcia ograniczają się nawzajem (Varela, 1996/2010, s. 61).

Rola fenomenologii nie ma więc sprowadzać się tylko do opisu doświadczenia, ale także ma aktywnie współuczestniczyć w budowaniu teorii naukowej, np. dokonując walidacji wyników badań neurokognitywnych.

Metoda neurofenomenologii składa się z trzech filarów:

1. (NPh1) Dane pierwszoosobowe pochodzące z uważnego badania doświadczenia z użyciem konkretnych metod pierwszoosobowych.
2. (NPh2) Formalne modele i narzędzia analityczne pochodzące z teorii układów dynamicznych, ugruntowane w ucieleśniono-enaktywnym podejściu do poznania.
3. (NPh3) Neurofizjologiczne dane pochodzące z pomiarów globalnych, integracyjnych procesów w mózgu (Lutz & Thompson, 2017, s. 463).

W swojej metodzie neurofenomenologia, po pierwsze, podejmuje problem rzetelności raportów introspekcyjnych i proponuje zamiast prostej metody introspekcyjnej wprowadzenie oryginalnej metody pierwszoosobowej, w której eksperymentator razem z badanym uzgadniają kategorie opisowe najlepiej wyrażające strukturalne własności badanego doświadczenia, a następnie osoba badana opisuje tymi kategoriami swoje przeżycia. Opis doświadczenia nie powstaje więc na bazie narzuconych odgórnie pojęć i teorii, tylko na bazie terminów pasujących osobom badanym. Należy dodać, że osoby uczestniczące w badaniu są wcześniej szkolone w wybranych elementach metody fenomenologicznej polegającej na uproszczonej wersji redukcji fenomenologicznej. Takie szkolenie dotyczy między innymi tego, w jaki sposób zawieszają zdroworoządkowe i teoretyczne przekonania na temat istoty doświadczeń i stanów umysłowych. Kolejnym, trzecim na liście, składnikiem metody neurofenomenologicznej jest zapis aktywności mózgu osób badanych za pomocą EEG. Zapis ten prowadzony jest w trakcie badania, a następnie analizowany pod kątem korelatów z opisem

doświadczenia. Narzędziem, wymienionym na drugiej pozycji, które ma pomóc w ustaleniu relacji pomiędzy opisami pierwszoosobowymi a danymi trzecioosobowymi, jest wspomniana powyżej teoria układów dynamicznych. Założeniem, jakie leży u podstaw takiego wyboru, jest stwierdzenie, że zarówno zjawiska fenomenologiczne, jak i procesy neuronalne cechuje określona dynamika, którą teoria układów dynamicznych może uchwycić.

W pierwszym, i jak dotąd jednym z nielicznych, eksperymencie neurofenomenologicznym (Lutz & Thompson, 2017) badano stan gotowości percepcyjnej. Badanym osobom przedstawiono zaszumiony obraz, z którego stopniowo wyłaniał się kształt geometryczny. Badani mieli uchwycić moment, w którym uświadamiają sobie, że widzą kształt. Ustalonymi wcześniej kategoriami, które miały oddać różne stopnie gotowości do postrzegania, były: stała gotowość, gotowość fragmentaryczna oraz niegotowość. Następnie próbowano skorelować te trzy kategorie z zarejestrowanymi w badaniu wzorcami EEG. Badanie, które miało charakter pilotażowy, pokazało, że można zastosować elementy metody fenomenologicznej do poprawy jakości danych pierwszoosobowych. Nie udało się natomiast w pełni zrealizować drugiego elementu metody, czyli formalnych modeli dynamicznych. Połączenie kategorii opisowych i wzorców aktywności EEG opierało się na prostej korelacji w czasie. Nie pokazano, czy i w jaki sposób wybrane parametry EEG związane są z własnościami fenomenologicznych kategorii doświadczenia percepcyjnego. Nie przedstawiono również modelu dynamicznego, który łączyłby aspekt doświadczeniowy z neuronalnym.

Podsumowując, neurofenomenologia jest nieredukcjonistycznym projektem fenomenologii znaturalizowanej. Rolą fenomenologii jest tutaj dostarczanie wiarygodnych opisów badanego doświadczenia i identyfikacja jego własności strukturalnych. Co więcej, fenomenologia mogłaby odgrywać rolę w walidacji empirycznych badań świadomości. Jednak neurofenomenologia, dopóki nie rozwinie drugiego elementu swojej metody, czyli modeli dynamicznych pełniących funkcję integrującą obie domeny, pozostanie interesującą propozycją teoretyczną.

Wybrane propozycje metodologiczne: fenomenologia fazy wstępnej

Fenomenologia fazy wstępnej (*front-loaded phenomenology*) jest propozycją naturalizacji fenomenologii poprzez włączenie wybranych pojęć i teorii fenomenologicznych do praktyki eksperymentalnej (np. Gallagher, 2017). Taki sposób naturalizacji jest dość liberalny i nie wymaga szkolenia badaczy lub osób badanych w metodzie fenomenologicznej ani pełnej znajomości teorii fenomenologicznych. Zespoły badawcze miałyby korzystać z koncepcji, pojęć i rozróżnień fenomenologicznych we wczesnych fazach prowadzenia badań, mianowicie na etapie projektowania eksperymentu. Przykładowymi pojęciami fenomenologicznymi, które miały zastosowanie w badaniach empirycznych, są zaczerpnięte z filozofii Merleau-Ponty'ego i jej współczesnych kontynuacji schemat i obraz ciała, poczucie sprawstwa i poczucie własności ciała, czy też koncepcja świadomości czasu i pojęcie „intencjonalność” z fenomenologii Husserla. Zdaniem Gallaghera:

fenomenologia zastosowana na początkowym etapie badań może być rozwinięta zarówno w analizy czysto fenomenologiczne (jak u Husserla), jak i eksperymenty neurofenomenologiczne. Po drugie, naturalną konsekwencją fenomenologii fazy wstępnej jest to, że tak jak w modelu neurofenomenologicznym, fenomenologia staje się częścią podstawy analitycznej do interpretowania wyników, a nie jedynie częścią danych, które mają zostać przeanalizowane (Gallagher, 2017, s. 449).

Za przykład fenomenologicznie inspirowanego badania mogą posłużyć eksperymenty z poczuciem sprawstwa (*sense of agency*) (zob. Gallagher, 2017). W badaniach tych wykorzystano opracowaną przez Gallaghera fenomenologiczną koncepcję poczucia sprawstwa i wywoływano eksperymentalnie sytuacje, w których osoby badane doświadczały zaburzenia tego poczucia w odniesieniu do podejmowanego cielesnie działania (przesuwanie punktu na ekranie za pomocą joysticka). Jednocześnie skanowano aktywność mózgu badanych za pomocą fMRI. Wyniki sugerują, że postrzeganie działań, których sprawcą jest ktoś inny, skorelowane jest z aktywnością w prawej dolnej korze ciemieniowej, natomiast gdy działanie odczuwane było jako spowodowane przez samych badanych, wtedy aktywna była przednia części wyspy.

Oceniając projekt fenomenologii fazy wstępnej, należy stwierdzić, że przedkłada ona pragmatyzm badawczy nad żmudne rozważania metodologiczne. Zwolennik takiego podejścia mógłby rzec: jeśli współpraca fenomenologów i kognitywistów działa w laboratorium, to znaczy, że jest to dobra droga i róbmy tak dalej. Nie jest jednak do końca jasne, jaką rolę aplikowane do badań koncepcje fenomenologiczne mają odgrywać w całym procesie badawczym. Po pierwsze, można potraktować pojęcia fenomenologiczne podobnie do innych pojęć naukowych lub potocznych pojęć psychologicznych, które wpływają na konceptualizację badanego zjawiska i przygotowanie projektu eksperymentu. Jednak taka możliwość nie pokazuje, dlaczego teorie fenomenologiczne miałyby być lepsze i dlatego preferowane względem innych koncepcji. Drugą opcją, na którą wskazuje sam autor tej koncepcji, jest, że koncepcje fenomenologiczne mogą pomóc w formułowaniu hipotez badawczych, które następnie mogłyby zostać poddane testowaniu w badaniu empirycznym. Nie jest to jednak zgodne z omawianymi w literaturze zastosowaniami fenomenologii fazy wstępnej, w których pojęcia fenomenologiczne są raczej wykorzystywane do zdefiniowania *explanandum* niż do formułowania hipotez, jak miało to miejsce w omawianym badaniu poczucia sprawstwa. Odrębną kwestią wymagającą dłuższej dyskusji jest, czy twierdzenia fenomenologiczne są w ogóle testowalne empirycznie. Gallagher twierdzi, że jest to możliwe, a w ten sposób naturalizacja doprowadzi do „wzajemnego oświecenia” fenomenologii i kognitywistyki. Trzecia możliwość, którą zdaje się wskazywać przytoczony powyżej cytat, przypisuje fenomenologii rolę analityczną i interpretacji wyników badań. Jeśli tak, to miejsce pojęć fenomenologicznych powinno znajdować się raczej po stronie *explanansu*, czyli w części wyjaśniającej badane zjawisko zamiast je opisujące. Podsumowując, fenomenologia fazy wstępnej jest dość prostym i naturalnym sposobem integracji pojęć fenomenologicznych z empiryczną praktyką badawczą. Należałoby natomiast doprecyzować etap zastosowania pojęć fenomenologicznych w całym procesie.

Refleksja systematyczna z wnioskami i rekomendacjami

Oceniając obecny stan debaty na temat naturalizacji fenomenologii, można stwierdzić, że pomimo pojawienia się kilku interesujących pomysłów teoretycznych, które przełożyły się na wzrost zainteresowania fenomenologią wśród kognitywistów i neuronaukowców, a także pomimo prób aplikacji teorii fenomenologicznych w praktyce badawczej nauk kognitywnych, nie doszło do wyraźnego postępu w procesie naturalizacji. Dalej nie wiadomo, w jaki sposób metodologicznie połączyć fenomenologię i nauki kognitywne. Jednym z powodów takiej sytuacji może być niewystarczający namysł nad tym zagadnieniem z wykorzystaniem narzędzi dostarczanych przez filozofię nauki. Rozpatrzenia wymagałyby chociażby relacja fenomenologii do różnych typów wyjaśnień naukowych. Jest to o tyle ważne, że eksplanacyjny charakter naturalizacji jako integracji jest przecież wymieniony w jej definicji, natomiast proponowane rozwiązania temat ten omijają lub podejmują w sposób niewystarczający. W szczególności należy się przyjrzeć typom wyjaśnień stosowanym w naukach kognitywnych, jak np. wyjaśnieniom funkcjonalnym, dynamicznym, mechanistycznym, a także określić, do jakiego stopnia, jeśli w ogóle, fenomenologia jest z nimi zbieżna. Ponadto należałoby odpowiedzieć na pytanie, czy fenomenologia oferuje swoisty typ wyjaśnień, a jeśli tak, to na czym miałby on polegać (zob. szczegółowe omówienie tych zagadnień w: Pokropski, 2021).

Pytając o możliwości eksplanacyjne fenomenologii, można podać jej w wątpliwość. Fenomenologia, posługując się metodą redukcijną, abstrahuje od relacji przyczynowych występujących w świecie, natomiast w przypadku zjawisk umysłowych zastępuje je relacjami motywacji (np. Husserl, 1962). Stanowi to potencjalny problem dla projektu naturalizacji, który zakłada wykazanie ciągłości pomiędzy opisywanymi przez fenomenologię własnościami przeżywanego doświadczenia a przyczynowymi procesami opisywanymi przez nauki przyrodnicze. Rozwiązanie, lub jego brak, problemu przyczynowości umysłowej może zdecydować o powodzeniu naturalizacji. Fenomenologia nie może więc dostarczyć naukowego, w znaczeniu przyczynowego, wyjaśnienia. Może za to dostarczać narzędzi do pierwszoosobowego opisu i analizy takich zjawisk, czym przyczynia się do ich lepszego zrozumienia.

Przykładowo badania fenomenologiczne rozszerzyły nasze rozumienie procesów poznawczych o aspekt ucieleśnienia, metody i pojęcia fenomenologiczne stosowane były również do lepszego zrozumienia zaburzeń psychicznych takich jak depresja czy schizofrenia.

Rozwiązaniem, które stara się ominąć problem przyczynowości, jest naturalizacja fenomenologii rozumiana jako integracja eksplanacyjna zgodna z założeniami mechanistycznego modelu wyjaśniania (Pokropski, 2021). Jednym z wniosków płynących z dyskusji w obszarze filozofii kognitywistyki i neuronauki jest stwierdzenie, że najbardziej rozpowszechniony w tych dyscyplinach jest mechanistyczny model wyjaśniania (np. Craver, 2007). Według zwolenników mechanicyzmu wyjaśnienie jakiegoś zjawiska polega na opisaniu przyczynowego mechanizmu odpowiedzialnego za powstanie tego zjawiska. Wyjaśnianie zjawisk umysłowych polegałoby więc na poszukiwaniu mechanizmów neuronalnych, które są za nie odpowiedzialne. Wieloaspektowa natura zjawisk umysłowych sprawia, że aby ich wyjaśnienie się powiodło, potrzeba zintegrowania wiedzy z różnych obszarów badań, np. z neurofizjologii, psychologii, modelowania komputerowego. Co ważne, mechanicyzm proponuje sposób takiej integracji. Zgodnie z tą koncepcją obszary badawcze, wskazując na własności hipotetycznych mechanizmów, nakładają ograniczenia na przestrzeń możliwych mechanizmów i w ten sposób przyczyniają się do charakterystyki tego właściwego i poszukiwanego mechanizmu (zob. np. Craver, 2007). Mechanistyczny model wyjaśniania, mimo licznych zalet, posiada również ograniczenia, m.in. w wyjaśnianiu złożonych zjawisk umysłowych, jak np. świadomość i jej zaburzenia, wymagających zastosowania metod pierwszoosobowych. Pomóc w tym mogłaby fenomenologia, która sama nie posiadając mocy eksplanacyjnej, może dostarczyć opisów i analiz struktury doświadczenia, a te z kolei mogą się przełożyć na sformułowanie ograniczeń eksplanacyjnych. Z tego względu integracja podejść fenomenologicznego i mechanistycznego wydaje się ciekawą perspektywą na naturalizację.

Projekt naturalizacji fenomenologii pozostaje otwarty. Jeśli sformułowana zostanie satysfakcjonująca propozycja fenomenologii znaturalizowanej, to mogłaby stanowić duży krok w budowaniu nauki o świadomości. Jednak niezależnie od powodzenia tego projektu wartościowe wydają się już same rozważania nad relacją między fenomenologią a naukami przyrodniczymi. Zrozumienie tej relacji jest ważne zarówno

dla fenomenologii, która raz to występuje w roli przewodniej i fundującej naukową racjonalność, a innym razem w roli dyscypliny partnerskiej, jak i dla nauk przyrodniczych wyrastających z nowożytnego matematycznego przyrodoznawstwa, dla których zjawiska takie jak świadomość pozostają zagadką.

BIBLIOGRAFIA

- Chalmers, D. (1996/2014). *Świadomy umysł: w poszukiwaniu teorii fundamentalnej*. Przeł. M. Miłkowski. Warszawa: Wydawnictwo Naukowe PWN.
- Craver, C.F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.
- Dreyfus, H., & Hall, H. (1982). Introduction. W: H.L. Dreyfus, & H. Hall (Red.), *Husserl, intentionality and cognitive science* (s. 1–27). Cambridge: MIT Press.
- Gallagher, S. (2017) Fenomenologia i projektowanie eksperymentów. Ku fenomenologicznie oświeconym naukom eksperymentalnym. Przeł. P. Nowakowski. W: J. Migasiński, & M. Pokropski (Red.), *Główne problemy współczesnej fenomenologii* (s. 440–457). Warszawa: Wyd. UW.
- Gallagher, S. (2018). Rethinking nature: Phenomenology and a non-reductionist cognitive science. *Australasian Philosophical Review*, 2(2), s. 125–137.
- Husserl, E. (1962). *Phänomenologische Psychologie: Vorlesungen Sommersemester 1925*, The Hague: Nijhoff.
- Husserl, E. (1974). *Idee czystej fenomenologii i fenomenologicznej filozofii*, księga II. Przeł. D. Gierulanka. Warszawa: PWN.
- Husserl, E. (1989). *Wykłady z fenomenologii wewnętrznej świadomości czasu*. Przeł. J. Sidorek. Warszawa: PWN.
- Lutz, A., & Thompson, E. (2017) Neurofenomenologia: Integrowanie doświadczenia subiektywnego i dynamiki neuronalnej w neuronauce świadomości. Przeł. L. Ciechanowski. W: J. Migasiński, & M. Pokropski (Red.), *Główne problemy współczesnej fenomenologii* (s. 458–486). Warszawa: Wyd. UW.
- Merleau-Ponty, M. (2001). *Fenomenologia percepcji*. Przeł. M. Kowalska, & J. Migasiński. Warszawa: Fundacja Aletheia.
- Moran, D. (2013). 'Let's look at it objectively': Why phenomenology cannot be naturalized. *Royal Institute of Philosophy Supplements*, 72, 89–115.

- Pokropski, M. (2021). *Mechanisms and consciousness: integrating phenomenology with cognitive science*. New York: Routledge.
- Roy, J.-M., Petitot, J., Pachoud, B., & Varela, F.J. (1999). Beyond the gap: An introduction to naturalizing phenomenology. W: J. Petitot, F.J. Varela, B. Pachoud, & J.-M. Roy (Red.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (s. 1–80). Stanford: Stanford University Press.
- Varela, F. (1996/2010). Neurofenomenologia: metodologiczne lekarstwo na trudny problem. Przeł. R. Poczobut. *Avant*, 1(210), 31–75.
- Zahavi, D. (2010). Naturalized phenomenology. W: S. Gallagher, & D. Schmicking (Red.), *Handbook of phenomenology and cognitive science* (s. 3–19). Dordrecht: Springer.

Jacek Bielas

Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-9771-4162>

<https://doi.org/10.35765/slowniki.457>

Ucieleśnione poznanie

Streszczenie

DEFINICJA POJĘCIA: Pojęcie ucieleśnionego poznania dotyczy istotnego znaczenia różnych aspektów całokształtu ciała, nie tylko układu nerwowego, w ścisłym związku ze środowiskiem jako konstytutywnego elementu systemu poznawczego.

ANALIZA HISTORYCZNA POJĘCIA: Geneza współczesnych badań nad ucieleśnieniem poznania wiąże się głównie z krytyką klasycznego ujęcia procesów poznawczych w kategoriach komputacjonizmu, według którego procesy te dzieją się zasadniczo w mózgu i polegają na obliczeniowym przetwarzaniu informacji w formie symbolicznych reprezentacji przedmiotu poznania. W toku tej krytyki pojęcie ucieleśnionego poznania stało się wiodącą ideą interdyscyplinarnego projektu badawczego określanego mianem ucieleśnionej kognitywistyki.

UJĘCIE PROBLEMOWE POJĘCIA: Pojęcie ucieleśnionego poznania, pomimo swego naukowego statusu, jawi się jako wieloznaczne, a projekt ucieleśnionej kognitywistyki jako mało spójny. Przede wszystkim jednak kwestia faktycznej roli cielesności w aktywności poznawczej również pozostaje nierozstrzygnięta w ramach spektrum stanowisk od różnych wersji komputacjonizmu, na jednym biegunie, poprzez „minimalne”, aż po „ekstremalne” ucieleśnienie poznania – na biegunie przeciwnym.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Współczesna problematyka ucieleśnionego poznania obejmuje szereg podejść do tej kwestii, które można różnie klasyfikować. Ta różnorodność przedmiotowa i metodologiczna wiąże się z wieloma trudnościami,

na przykład w zakresie replikacji badań. Jednocześnie natomiast rozwój projektu ucieleśnionej kognitywistyki wydaje się sprzyjać zarówno poszerzeniu zakresu zagadnień, jak i właśnie zachowaniu spójności tych zagadnień, które podejmowane są w dziedzinie podejmowanych w dziedzinie nauk o poznaniu, co wymaga jednak dalszych wysiłków o charakterze teoretycznym oraz empirycznym.

Słowa kluczowe: procesy poznawcze, cielesność/ciało,
komputacjonizm, ucieleśnione poznanie, 4 E

Definicja pojęcia

W literaturze przedmiotu pojęcie ucieleśnionego poznania (*embodied cognition*) wiąże się z poglądem, zgodnie z którym aktywność poznawcza organizmów żywych, w tym również człowieka, oraz robotów warunkowana jest funkcjonowaniem całokształtu ich budowy, a nie tylko jej fragmentów (takich jak mózg czy procesor). Zagadnienie ucieleśnionego poznania podejmowane jest w ramach ujęć podkreślających rolę różnych aspektów cielesności w procesach poznawczych, ale szczególnie wagę przywiązuje się pod tym względem do sensomotorycznych właściwości ciała działającego, aby mogło osiągnąć określone cele w otaczającym je środowisku. W ten sposób, na przykład, orientacja utajonej uwagi wzrokowej polegałaby na przygotowaniu (*motor preparation*) i osiągnięciu gotowości do ruchu (*motor readiness*) gałki ocznej w kierunku określonego miejsca i/lub obiektu w przestrzeni, emocja strachu byłaby niczym innym jak tylko doświadczeniem przygotowania (się) i gotowości organów wewnętrznych i mięśni do ucieczki przed niedźwiedziem, a „ja” sprowadzałoby się do doświadczenia siebie jako rozciągniętego, ale ograniczonego w przestrzeni, a przy tym konfigurującego odpowiednio schemat swojego ciała, aby „sięgnąć” po obiekty w świecie i wykorzystać je do swoich celów. Aktywność poznawcza rozumiana w kategoriach cielesnej mobilności doświadczana jest w toku oporu i ciężenia ciała, które wynikają z jego fizykalnego charakteru, a poprzez nie oporu i ciężenia ze strony innych obiektów w świecie i dlatego wymaga siły oraz wysiłku.

Motoryczna aktywność cielesnej struktury w środowisku wymaga oczywiście również koordynacji jej elementów składowych w ramach wewnętrznej przestrzeni ciała w relacji z zewnętrzną przestrzenią otoczenia – do tego pod presją czasu. Centralną rolę w tej koordynacji odgrywa mózg, ale jednocześnie procesy neuronalne zdeterminowane są – siłą rzeczy – tym, co mają koordynować. Ujmowana w ten sposób cielesność pozostająca w ścisłym związku z otoczeniem stanowi element konstytutywny dla systemu poznawczego: ciało – mózg – środowisko. Przy czym kwestia dynamiki wzajemnych relacji pomiędzy składowymi tego systemu wydaje się pozostawać sprawą otwartą.

Uważa się, że sensomotoryczny i utylitarny charakter poznania może wynikać z uwarunkowań ewolucyjnych (Wilson, 2008). W rozwoju

filogenetycznym człowiek pojawił się relatywnie niedawno, a przedtem główną rolą organizmów żywych było koordynowanie ruchów swojego ciała, aby osiągnąć w przestrzeni konkretny skutek polegający na zbliżeniu się i wejściu w kontakt z określonymi obiektami lub oddaleniu od nich. Choć trudno byłoby odmówić człowiekowi wyjątkowych zdolności do świadomej refleksji, długofalowego planowania, operowania pojęciami abstrakcyjnymi i myślenia pozakontekstowego, to jednak poznawczy walor i specyfika pojęcia ucieleśnionego poznania jako elementu podejścia ewolucyjnego miałyby polegać na jego użyteczności w wyjaśnieniu rozwoju ludzkich zdolności poznawczych na podłożu zdolności sensomotorycznych.

Problematyka ucieleśnionego poznania podejmowana jest z osobna lub na sposób interdyscyplinarny na gruncie filozofii, językoznawstwa, psychologii, neuronauki, socjologii, informatyki, robotyki itd. W ostatnich dziesięcioleciach nastąpiła intensyfikacja rozważań na ten temat i pojęcie ucieleśnionego poznania utorało sobie drogę do głównego nurtu badań nad poznaniem, czy w ogóle funkcjonowaniem organizmów żywych i maszyn, stanowiąc wiodącą ideę interdyscyplinarnego programu badawczego określanego mianem ucieleśnionej kognitywistyki (*embodied cognitive science*). Trudno jednak oprzeć się wrażeniu, iż zarówno teoretyczne podejścia do tej sprawy, jak i propozycje jej rozstrzygnięcia nie są w zasadzie ani niczym szczególnie nowym, ani przynajmniej oryginalnym w historii myśli europejskiej, a ich początków można dopatrywać się w jakiejś mierze już u jej zarania. Z drugiej strony jednakże warto również podkreślić aktualny rozwój metod (np. eksperymentalnych) weryfikacji różnych ujęć zagadnienia ucieleśnionego poznania – głównie na gruncie nauk empirycznych – a także pojawienie się pewnych nowych okoliczności i kontekstów, w których problematyka ta może być podjęta jeszcze w inny sposób niż dotychczas. Okazję taką dają, na przykład, rozwój algorytmów inteligencji obliczeniowej (w tym awatary i chatboty) czy robotyka.

Analiza historyczna pojęcia

Antykartezjański rodowód pojęcia ucieleśnionego poznania

Chociaż historia zagadnienia statusu ciała w całokształcie funkcjonowania poznawczego wpisuje się już w nurt platońskiego idealizmu i arystotelesowskiego hylemorfizmu, to rodowodu nowożytnej refleksji na ten temat upatruje się zazwyczaj w kartezjańskim dualizmie, a ściślej rzecz biorąc – w jego krytyce. Według Kartezjusza ciało ludzkie to rozciągły obiekt materialny, którego istnienie dane jest psychicznemu, nierozciąglemu podmiotowi jedynie jako przedmiot występujący zawsze w tym samym miejscu i czasie, ale niejako obok. W *Uwagach skierowanych przeciwko pewnemu programowi* Kartezjusz stawia również tezę, że ciało przekazuje umysłowi poprzez jego organy zmysłowe bodźce ze świata zewnętrznego, które nie są ideami, lecz które umysł zamienia na idee, co może dawać asumpt do przypuszczeń o „znakowym” charakterze wrażeń stanowiących „rekonstrukcję sygnałów przesyłanych przez aparat nerwowy” (za: Czerniak & Michalski, 2008, s. 14). Na gruncie współczesnej kognitywistyki idee, o których tu mowa, jako „znaki” będące wynikiem przetwarzania przez ciało bodźców można by zinterpretować w kategoriach reprezentacji, a proces przetwarzania bodźców jako posiadający charakter obliczeniowy, chociaż sam Kartezjusz oczywiście nie dostarcza jeszcze podstaw dla tego typu rozstrzygnięć.

Na przełomie XVIII i XIX wieku w pracach innego francuskiego myśliciela, Maine’a de Birana, pojawia się ujęcie przełamujące sztywne granice kartezjańskiego dualizmu, w ramach którego związek *res cogitans* (duszy, umysłu, intelektu) i ciała jako *res extensa* ma w zasadzie jedynie charakter przyczynowo-skutkowy. Maine de Biran zauważa natomiast, że w aktach bezpośredniej apercpcji dana jest w e w n ę t r z n a p r e s t r z e ń w ł a s n e g o c i a ł a , w k t ó r e j g r a n i c a c h r o z p o ǫ s c i e r a s i ę j a s t a n o w i ą c e s i ł ę w y r a ż a j ą c ą s i ę w w y s i ł k u i r u c h u . Ó w w y s i ł e k u ś w i a d a m i a n y j e s t j a k o j a , k t ó r e „ d ą ż y d o r u c h u ” , a l e j e d n o c z e ś n i e d o ś w i a d c z a o p o r u c i a ł a , a p o p r z e z n i e o p o r u z e s t r o n y i n n y c h r z e c z y w ś w i e c i e (M a i n e d e B i r a n , 1 9 7 0 , s . 5 5) . C i a ł o n i e j e s t j u ż w i ę c r o z u m i a n e w y ł ą c z n i e j a k o p o z b a w i o n e ż y c i a p r z e d ł u ż e n i e u m y ś l u , f a k t f i z j o l o g i c z n y i p r z e d m i o t

w świecie, lecz poruszane jest niejako wewnątrznie i doświadczane wisceralnie i kinestetycznie.

Cielesność w fenomenologii

W przedstawionym przez Maine'a de Birana ujęciu tego, jak to jest być cielesnym, przynajmniej na poziomie opisu, łatwo można dopatrzeć się analogii do intencjonalnej struktury świadomości zaprezentowanej później na gruncie fenomenologii. Jej wiodący przedstawiciele (Husserl, Heidegger, Merleau-Ponty) podkreślali rolę cielesności w całokształcie doświadczenia, analizując różne sposoby warunkowania przez ciało – w tym również jego fizyczne aspekty – procesów poznawczych w aktach o charakterze temporalnym, przestrzennym, uwagowym, kinestetycznym, społecznym. W pracach Edmunda Husserla (1975) pojawiło się rozróżnienie na ciało jako bryłę rozumiane w kategoriach fizycznych i biologicznych (*Körper*), dane w perspektywie trzeciej osoby, oraz żywe ciało (*Leib*), przeżywane w perspektywie pierwszej osoby.

Za autora najbardziej zaawansowanego ujęcia zagadnienia ciała rozumianego nie tylko w aspekcie przestrzennym, ale także przeżyciowym, uważa się natomiast Maurice'a Merleau-Ponty'ego (2001). W centrum charakterystyki statusu cielesności w całokształcie aktywności poznawczej Merleau-Ponty sytuuje pojęcie schematu ciała, które obejmuje przestrzenną i czasową orientację w świecie i jednocześnie motoryczną funkcję podejmowania tam konkretnych, praktycznych zadań. Merleau-Ponty rozwija swoją koncepcję cielesności, opierając ją na fundamentalnym pojęciu „prerefleksyjnej intencjonalności”, która jako przestrzenna motoryczność konstytuuje wszystkie formy aktywności człowieka, czyli ciała-podmiotu, kierującego się ku konkretnym rzeczom w otoczeniu, aby ich doświadczyć, zrozumieć je i poddać manipulacji. Przestrzeń ciała i przestrzeń zewnętrzną tworzą w ten sposób nierozdzielny, praktyczny system, w którym ciało dostarcza wspólnego podłoża umożliwiającego poznawanie obiektów w przestrzeni jako celów do działania. Początkowe kartezjańskie „ja myślę” przybiera zatem postać cielesnego „ja mogę”, które zawsze znajduje się w świecie.

Propozycja połączenia fizycznej i biologicznej charakterystyki ciała, w perspektywie trzeciej osoby, z jednoczesnym wglądem w prywatną

sferę jego doświadczania, w perspektywie pierwszej osoby, przedstawiona została przez Francisca Varełę (1996/2010) w kategoriach projektu neurofenomenologii. Propozycja ta miała pierwotnie stanowić remedium na trudny problem świadomości formułowany w kategoriach luki eksplanacyjnej (Levine, 1983) dotyczącej rozbieżności pomiędzy fenomenalnym charakterem doświadczenia a fizyczną i biologiczną naturą ciała (mózgu). Bez względu na ocenę efektywności neurofenomenologii jako projektu naturalizacji fenomenologii warto podkreślić jego wpływ na rozwój i praktyczne wykorzystanie metod łączących podejście fenomenologiczne z psychologią eksperymentalną i neurologią. Przykładem takiego połączenia może być metoda neurofeedbacku (patrz np. Petitmengin, Navarro & Le Van Quyen, 2007). Poprzez umożliwienie podmiotowi wglądu w swoje procesy neuronalne jako korelatu stanów mentalnych metoda ta miałaby rozwijać zdolności samokontroli procesów psychologicznych jednostki. Warto również przy tej okazji zauważyć, że techniki kontroli stanów psychicznych poprzez wpływ na stany cielesne znane są już i praktykowane od tysięcy lat, chociażby w systemie jogi.

Krytyka komputacjonizmu jako współczesny kontekst genezy pojęcia ucieleśnionego poznania

Mniej więcej od połowy XX wieku, wraz z pracami Alana Turinga nad projektem sztucznej inteligencji, w podejściu do fenomenu poznania zaczęła dominować paradygmat, według którego działanie umysłu, podobnie jak komputera, sprowadza się do arytmetycznych operacji na symbolach reprezentujących cechy świata zewnętrznego. Ze względu na ujęcie działania umysłu poprzez analogię z pracą komputera paradygmat ten zaczęto określać mianem komputacjonizmu. Cechą wspólną podejść opartych na idei komputacjonizmu jest opis poznania jako procesu fazowego składającego się z następujących po sobie etapów: transdukcji energii bodźca na jej symboliczny odpowiednik oraz przetwarzania symboli zgodnie z określonymi regułami, co prowadzi do określonego wyniku w postaci gramatycznych struktur językowych, wyodrębnienia ciągu słów pośród innych, rozwiązania problemu logicznego, rozpoznania bodźca w odniesieniu do składowanych w pamięci wzorców lub

trójwymiarowego obrazu świata. Wszystkie te operacje mają odbywać się w układzie nerwowym organizmu, który umożliwia kodowanie bodźców w formie języka myśli tak, jak to się dzieje w przypadku języków programowania komputerowego. W ten sposób w ramach komputacjonizmu, traktowanego jako dominujący nurt w naukach kognitywnych, umysł, czyli w gruncie rzeczy mózg, opisywany jest w kategoriach biologicznego komputera, który na sposób algorytmiczny operuje symbolami reprezentującymi elementy świata zewnętrznego.

Okolo połowy XX wieku ujęcie procesów poznawczych w kategoriach komputacyjnych zdominowało obszar nauk o poznaniu również w odniesieniu do takich zagadnień jak na przykład: teoria języka, uwaga, percepcja wzrokowa czy teoria rozwiązywania problemów. Jednocześnie jednak zwracano uwagę na szereg ograniczeń w wyjaśnianiu poznania w kategoriach komputacjonizmu, na przykład w odniesieniu do pytania, jak procesy algorytmiczne przebiegające w mózgu miałyby prowadzić do świadomych i jakościowo różnych doświadczeń lub, jako procesy wewnętrzne organizmu, reprezentować stany świata zewnętrznego. Istotnym elementem tej krytyki było również zbyt wąskie ujmowanie – zarówno w ramach klasycznego komputacjonizmu, jak i jego wersji w postaci koneksjonizmu – granic systemu poznawczego. Krótko mówiąc, krytyka ta odnosiła się, i nadal zresztą się odnosi, do „mózgocentryzmu” w podejściu do poznania jako dziejącego się wyłącznie „w głowie”.

Psychologia ekologiczna

Zgoła inne podejście do poznania, chociaż nawiązujące w znacznej mierze do pojęć przedstawionych wcześniej na gruncie fenomenologii przez Maurice’a Merleau-Ponty’ego („prerefleksyjna intencjonalność”) i Martina Heideggera (np. „poręczność”), zaproponował James Gibson (1966) w ramach psychologii ekologicznej. Jej podstawowym założeniem stało się ujęcie cielesnej struktury i jej motorycznych możliwości w zakresie orientacji i działania w świecie jako fundamentu dla procesów poznawczych u człowieka i zwierząt. Gibson koncentruje się tu na badaniu percepcji wzrokowej. Wyjaśnia ją w kategoriach bezpośredniej percepcji otoczenia, która nie jest zapośredniczona przez inne procesy

poznawcze, lecz zależy od usytuowania organizmu w środowisku i jego aktywności motorycznej. Jest to propozycja rozwiązania zagadnienia zbyt małej liczby bodźców docierających do narządów zmysłów w jednostce czasu, aby system poznawczy mógł je odpowiednio zinterpretować. Zgodnie z podejściem komputacyjnym informacje te są przetwarzane w mózgu w relacji do składowanych tam już wcześniej struktur poznawczych na zasadzie integracji procesów oddolnych (*bottom up*) i odgórných (*top down*). Według psychologów ekologicznych natomiast adekwatna interpretacja bodźców odbywa się w toku nieustannego ruchu całego ciała organizmu i scalania elementów poznawanego przedmiotu w jego całokształcie. Tak rozumiana percepcja nie polega zatem na odbiorze bodźców z otoczenia, które stają się następnie danymi podlegającymi operacjom obliczeniowym posiadającym charakter procesów mózgowych przebiegających wyłącznie w głowie, ale jest procesem realizowanym przez całe ciało działające w środowisku. *Nota bene* ujęcie procesów poznawczych jako ściśle uwarunkowanych motoryką spotyka się współcześnie również na gruncie eksperymentalnych badań nad, między innymi, uwagą wzrokowo-przestrzenną rozumianą w kategoriach działania dla percepcji (*action for perception*). Craighero i in. (1999) pokazali na przykład, że efektywność percepcji przedmiotów zależy od przygotowania ruchu dłoni poprzez jej odpowiednie ułożenie, aby je uchwycić.

Istotnym elementem psychologii ekologicznej stało się również pojęcie afordancji (*affordances*), które odnosi się do możliwości zachowań organizmu warunkowanych jego materialną konstytucją (np. stan skupienia czy kształt ciała). W ten sposób całokształt psychologicznego funkcjonowania, w tym także procesy percepcyjne oraz znaczenie pojęć, u człowieka i zwierząt warunkowane są we wzajemnych relacjach pomiędzy fizycznymi właściwościami rzeczy z jednej strony oraz fizykalnymi właściwościami i motorycznymi możliwościami wynikającymi ze struktury ciała organizmu z drugiej strony. Znaczenie obiektów w świecie zależy zatem od tego, w jaki sposób można nawiązać z nimi fizyczny kontakt i co można z nimi zrobić w zależności od cielesnych możliwości podmiotu. A zatem percepcja i pojęcie, na przykład, parasola nie zależałyby wyłącznie od możliwości obliczeniowych w procesie przetwarzania informacji na temat tego przedmiotu, zachodzących tylko w głowie, lecz od tego, co można z nim zrobić w zależności od interakcji

pomiędzy wodą a skórą człowieka oraz faktem posiadania przez niego dłoni o określonej budowie. Z tego względu parasol może istnieć jako coś zupełnie innego dla innych organizmów (np. nietoperza czy ryby) lub w ogólnie nie być przez nie spostrzeganym.

Cielesność sztucznej inteligencji w systemach robotycznych

W dziedzinie rozwoju systemów robotycznych dominował początkowo pogląd, zgodnie z którym procesy poznawcze sprowadzają się właściwie do operacji wykonywanych na abstrakcyjnych symbolach w centralnych modułach tych urządzeń. Przykładem takiego podejścia był jeden z pierwszych robotów o nazwie Shakey skonstruowany w Artificial Intelligence Laboratory w Stanford. Jego zadaniem było poruszanie się w pomieszczeniu i omijanie niektórych przedmiotów oraz przesuwanie innych. Zadania te wykonywane były dzięki programowi STRIPS funkcjonującemu na zasadzie sekwencji: rejestracja-modelowanie-planowanie-działanie (*sense-model-plan-act*). Kamera rejestruje elementy otoczenia, komputer przekształca je na symboliczne, czyli abstrakcyjne, modele, program porównuje je z uprzednio zdefiniowanymi modelami przechowywanymi w pamięci komputera i uruchamia zaplanowany również wcześniej program działania urządzenia. Tego typu mechanizm działania robotów, oparty na zasadzie tworzenia reprezentacji, uznany został za mało efektywny, a wręcz, być może, nieadekwatny. W artykule *Intelligence without representation* Rodney Brooks (1991) poddał krytyce stosowany w dziedzinie badań nad rozwojem sztucznej inteligencji i urządzeń robotycznych model procesów poznawczych oparty na zasadzie reprezentacji. Zgodnie z tym modelem centralny moduł systemu dokonuje operacji obliczeniowych na danych reprezentujących elementy otoczenia po ich przekształceniu na podstawie danych wejściowych (*input*) trafiających do systemu za pośrednictwem modułów zewnętrznych. Taki proces jest nieefektywny, ponieważ wymaga dodatkowych zasobów na tworzenie reprezentacji. Shakey na przykład był bardzo wolny. Brooks uważa, że zamiast tworzyć reprezentacje „lepiej jest wykorzystać świat jako swój własny model” (1991, s. 139, przeł. J.B.) i wystarczy, gdy system

będzie czerpał informacje bezpośrednio ze swojego otoczenia, potrafiąc dzięki temu dynamicznie adaptować się do zmian, które w nim zachodzą. Zamiast mechanizmu komputacji w projektowaniu i budowie swoich „Stworzeń” (*Creatures*) Brooks wykorzystuje podejście oparte na zasadach sformułowanych na gruncie psychologii ekologicznej Gibsona. Jego roboty funkcjonują na bazie rozbudowanego systemu czujników rejestrujących otoczenie i połączonych bezpośrednio z mechanizmami wykonywania określonych ruchów przez system efektorów bez pośrednich etapów tworzenia reprezentacji środowiska. W tak działającym urządzeniu kluczowa zatem staje się jego budowa i wynikające z niej fizyczne zdolności, które uzasadniają oprogramowanie służące obsłudze aktywności ruchowych. Brooks twierdzi, że system reprezentacji jest tu zbędny i wystarczy bezpośrednie połączenie „percepcji z działaniem” (*perception to action*) – *notabene* określając często ten drugi człon mianem zachowania (*behavior*). W powyższym opisie i podejściu do zasad funkcjonowania urządzeń robotycznych trudno nie zauważyć analogii do behawioryzmu na gruncie psychologii i formuły *S – R* (*bodziec – reakcja*). Jednakże pomimo niewątpliwej efektywności robotów konstruowanych przez Brooksa w porównaniu do poprzednich maszyn działających w sposób komputacyjny oraz wpływu jego koncepcji na tę gałąź przemysłu zagadnienie domniemanego braku systemu reprezentacji w proponowanych przez niego rozwiązaniach budzi wątpliwości. Najogólniej rzecz ujmując, zwraca się pod tym względem uwagę na konieczność istnienia oprogramowania służącego do selekcji informacji z otoczenia oraz wyboru i uruchamiania określonej akcji, które to oprogramowanie może posiadać cechy reprezentacji w zależności od jej przyjętego rozumienia.

Ujęcie problemowe pojęcia

Początek obecnej popularności idei ucieleśnionego poznania w naukach kognitywnych dość powszechnie wiązany jest z pojawieniem się pracy *The Embodied Mind* Francisca Vareli, Eleanor Rosch i Evana Thompsona (1993), w której autorzy wystąpili przeciwko formalnemu ujęciu procesów poznawczych jako realizowanych wyłącznie w mózgu liczbowych operacji na symbolicznych reprezentacjach. Zamiast tego

zapropowali wielowymiarową analizę umysłu, i w konsekwencji procesów poznawczych, obejmującą nie tylko mózg, ale także poza-neuronalne aspekty ciała (fizyczne, biologiczne w szerokim znaczeniu, fenomenologiczne) pozostającego w ścisłym związku z otaczającym je środowiskiem. Propozycja autorów w pełni przyczyniła się do powstania projektu ucieleśnionej kognitywistyki (*embodied cognitive science*) opartego na pojęciu poznania ucieleśnionego. Jego zwolennicy koncentrują się więc na ciele jako takim i jego całokształcie jako istotnym czynniku poznania, które również rozumiane jest w jak najszerszym sensie (Dąbrowski, 2021).

Problematyka ucieleśnionego poznania w metateoretycznym związku z przeciwstawianymi jej ujęciami o charakterze komputacjonistycznym obejmuje szereg zagadnień składowych. Do zagadnień takich można zaliczyć kwestię proporcji w udziale czynników komputacyjnych i cielesnych w całokształcie poznania. Chodziłoby tu również o dynamikę tych proporcji na różnych etapach procesu poznawczego. Czy w pierwszej fazie dominują, na przykład, czynniki czysto obliczeniowe, a ciało przejmuje wiodącą rolę dopiero w późniejszych etapach? Czy w związku z tym definiowanie parametrów ruchu dla takich efektorów jak gałka oczna czy dłoń na premotorycznym etapie orientacji uwagi można określić mianem programowania ich aktywności o charakterze czysto komputacyjnym, czy może etap ten nosi już cechy ucieleśnienia? Jeżeli tak, to na czym takie ucieleśnienie miałyby polegać?

Współcześnie wielu badaczy zgadza się co do tego, że ucieleśnienie stanowi warunek *sine qua non* każdej formy inteligencji, naturalnej czy sztucznej (Ziemke, 2015). Trudno byłoby na przykład mówić o istnieniu inteligencji wyłącznie w formie jakiegoś abstrakcyjnego algorytmu bez jego fizycznej (cielesnej) konkretyzacji. Tym niemniej kwestią otwartą pozostaje pytanie, co to miałyby znaczyć, że poznanie jest ucieleśnione, co wiąże się oczywiście z pytaniem, czym w ogóle jest i jakiego typu może być ciało. Uważa się na przykład, że ludzkie procesy poznawcze miałyby być specyficzne właśnie ze względu na charakterystykę realizującego je cielesnego podmiotu. Można by więc zapytać, w jakiego typu ciało należałoby wyposażyć systemy AI jako model inteligencji naturalnej. Warto przy tej okazji również podkreślić, iż pomimo popularności projektu ucieleśnionej kognitywistyki pogląd, że inteligencja wymaga ciała, nie jest podzielany przez wszystkich (Ziemke, 2015).

A zatem sama kwestia faktycznego statusu cielesności dla całokształtu procesów poznawczych pozostaje otwarta w ramach spektrum stanowisk od różnych wersji komputacjonizmu począwszy, poprzez „minimalne”, aż po „ekstremalne” ucieleśnienie poznania. Oznacza to więc, że obliczeniowe teorie umysłu nadal brane są pod uwagę. Konkretnym przykładem takiego stanu rzeczy jest toczący się obecnie spór na gruncie psychologii eksperymentalnej dotyczący natury uwagi wzrokowej pomiędzy zwolennikami tzw. premotorycznej teorii uwagi a badaczami stojącymi na stanowisku, że jest ona fenomenem niezależnym od procesów motorycznych, przynajmniej w przypadku endogennej uwagi wzrokowej. Należy jednak również przy tej okazji podkreślić różnicę kategoriałną pomiędzy tezami o umyśle a tezami o poznaniu.

Oczywiście zagadnieniem zasadniczym również w kontekście pytania, czy poznanie w ogóle musi, a jeżeli tak, to jak bardzo może, być ucieleśnione, jest rozumienie samego ciała/ucieleśnienia. Ziemke (2015) na przykład przytacza sześć takich rozumień funkcjonujących na gruncie literatury przedmiotu, podkreślając jednocześnie trudności w ich ścisłym zdefiniowaniu: 1) „sprzężenie strukturalne” pomiędzy podmiotem a środowiskiem, 2) „ucieleśnienie historyczne” wynikające z historii sprzężenia strukturalnego, 3) „ucieleśnienie fizyczne”, 4) „ucieleśnienie organizmoidalne” dotyczące organizmopodobnych form cielesnych, takich jak na przykład roboty humanoidalne, 5) „ucieleśnienie organizmowe” żywych systemów (autopojetycznych) i 6) „ucieleśnienie społeczne”.

Zapewne dla uzyskania pełnej charakterystyki cielesności należy uwzględnić również jej aspekt fenomenalny, na co zwraca uwagę w prezentowanym przez siebie zestawieniu Dąbrowski (2021). Chodzi mianowicie o obraz ciała, schemat ciała, który dany raczej jedynie w postaci poczucia, oraz doznania czysto somatyczne i te, które posiadają charakter emocjonalny, a zatem i motywacyjny.

Te różne ujęcia cielesności mogą zostać zidentyfikowane jako konstytuujące różne rozumienia i podejścia do problematyki poznania jako poznania ucieleśnionego.

Refleksja systematyczna z wnioskami i rekomendacjami

Współczesne koncepcje ucieleśnionego poznania: 4 E

Różne rozumienia i aspekty cielesności stanowią więc podstawę czterech podejść do problematyki poznania, które wydają się odgrywać wiodącą rolę na gruncie ucieleśnionej kognitywistyki. Określa się je zbiorczo terminem „4 E” od pierwszej litery angielskich przymiotników charakteryzujących poznanie: enaktywne lub skorelowane z działaniem (*enactive*), zakorzenione w świecie (*embedded*), rozszerzone (*extended*) oraz ucieleśnione (*embodied*) – stanowiące na tym obszarze badań pojęcie najbardziej ogólne i w zasadzie obejmujące je wszystkie.

Enaktywizm

Enaktywizm to podejście do poznania jako konstytuowanego w toku aktywności sensomotorycznej. W ten sposób poznanie stanowi emergentny proces warunkowany działaniem układu wzajemnie powiązanych elementów: mózg – ciało – środowisko. Podejście to zostało początkowo przedstawione przez Varełę, Rosch i Thompsona (1993) w wersji, którą można określić mianem „enaktywizmu autopoietycznego”. Autorzy zrównują w niej zasadniczo poznanie z życiem, podkreślając że wszystkie żywe organizmy (nawet tak pierwotne jak bakterie) wykazują w jakimś stopniu zdolności poznawcze, reorganizując swoje funkcjonowanie w toku dynamicznej interakcji z warunkami środowiska. Intencjonalność i doświadczeniowy charakter percepcji podkreślany jest szczególnie w wersji „enaktywizmu sensomotorycznego”. W ujęciu tym, podobnie zresztą jak w Fenomenologii percepcji Merleau-Ponty’ego, percepcja polega na motorycznej eksploracji środowiska, w wyniku której dochodzi do ustalenia wzorców zależności pomiędzy ruchami organizmu, mentalną treścią percepcji i światem. W ten sposób to, co spostrzegane, uwarunkowane jest tym, co organizm jest w stanie zrobić – a wręcz percepcja jest tym, co robimy. W wersji „enaktywizmu radykalnego” eliminacji ulega natomiast samo pojęcie treści mentalnych

rozumianych w kategoriach reprezentacji umysłowych jako nieefektywnych z adaptacyjnego punktu widzenia i zbędnego eksplanacyjnie.

Poznanie zakorzenione

Związek aktywności poznawczej ze środowiskiem przybiera jeszcze inną postać w perspektywie podejścia opartego na koncepcji poznania zakorzenionego w świecie. W ujęciu tym środowisko traktowane jest jako konstytutywna składowa samego procesu poznania rozumianego szerzej niż tylko procesy zachodzące w obrębie poznającego podmiotu. W tej perspektywie funkcjonowanie organizmów rozumianych w kategoriach systemów otwartych zależy ściśle od warunków środowiska, w którym się rozwijają i w którym funkcjonują. W ten sposób warunki te wpływają na ich budowę fizjologiczną, wykorzystanie wynikającego z niej potencjału i schematy działania. W przypadku człowieka zwraca się tutaj uwagę nie tylko na uwarunkowania przyrodnicze, ale również na rzeczywistość społeczną i kulturową.

Poznanie rozszerzone

W ramach koncepcji poznania rozszerzonego, którą od pozostałych ujęć odróżnia jej wyraźnie filozoficzny charakter, środowisko traktowane jest jako potencjalnie pomocne narzędzie służące do efektywniejszego radzenia sobie z zadaniami. Podejście to rozwinęło się na podstawie koncepcji poznania rozproszonego (Hutchins, 1996), w ramach której podkreśla się wymagający charakter procesów poznawczych przebiegających pod presją czasu. Może to prowadzić do konieczności przeniesienia ciężaru aktywności poznawczej na środowisko, które staje się w ten sposób częścią systemu poznawczego. Obsługiwanie skomplikowanych urządzeń, takich jak na przykład duży statek lub samolot, jest możliwe tylko wtedy, gdy czynność ta zostanie rozłożona na większą liczbę specjalistów. Również uporządkowanie książek na półce według pewnego klucza ułatwia później znalezienie tej potrzebnej. W takich przypadkach poznawcze możliwości podmiotu wzrastają w warunkach otoczenia zorganizowanego odpowiednio pod względem fizycznym lub

społecznym. Koncepcja poznania rozszerzonego, przedstawiona przez Andy'ego Clarka i Davida Chalmersa (1998/2008) w pracy *Umysł rozszerzony*, cechuje funkcjonalizm. Oparta jest na założeniu, według którego ta sama funkcja poznawcza może być realizowana na różne sposoby. Jeżeli jakieś działanie prowadzi do osiągnięcia określonego celu, to spełnia kryteria procesu poznawczego nawet wówczas, gdy jest realizowane, wykraczając poza granice mózgu. Dzieje się tak na przykład w przypadku narzędzi wspomagających możliwości ciała oraz odpowiedników treści umysłowych na nośnikach zewnętrznych. Umysł rozszerzony, rozumiany w kategoriach funkcjonalizmu, jest tutaj zatem systemem poznawczym, który przekracza granice dualizmu umysł – ciało.

Ucieleśnione poznanie

Pojęcie to stanowi niejako wspólny mianownik dla różnych stanowisk, które w reakcji na formalizm i mózgocentryzm w podejściu do poznania wyjaśnianego w kategoriach procesów obliczeniowego manipulowania symbolami dziejących się wyłącznie w głowie przypisują istotną rolę również pozamózgowym i pozaneuronalnym aspektom ciała w ścisłym związku ze środowiskiem jako konstytutywnym elementem systemu poznawczego. W ten sposób – jak ujmuje to Louise Barrett (2011) w swojej pracy *Beyond the Brain: How Body and Environment Shape Animal and Human Mind* – aktywność poznawcza organizmu

warunkowana jest (*reflect*) budową jego układu nerwowego, który warunkowany jest z kolei budową jego ciała, które kształtowane jest z kolei przez niszę środowiskową, którą ów organizm zajmuje (s. 11, przeł. J.B.).

Przy tym poprzez aktywność poznawczą należy tutaj rozumieć nie tylko jej aspekt behawioralny, ale również fenomenalny. Aktywność poznawcza człowieka różni się więc zasadniczo od aktywności poznawczej zwierząt, ponieważ zwierzęta posiadają inne ciała, inne układy nerwowe oraz zamieszkują inne habitaty i dlatego nawet jeżeli zachowania eksploracyjne jednych i drugich mogą wydawać się podobne, człowiek nie będzie w stanie zrozumieć, jak to jest poznawać świat, będąc na przykład nietoperzem.

Cztery wiodące podejścia do kwestii ucieleśnionego poznania (4 E) w ramach ucieleśnionej kognitywistyki obejmują zbiór zagadnień składowych, które podejmowane są tam w różny sposób i różnie się je rozstrzyga. Stąd zresztą pojawiające się również opinie o niejednorodności, niespójności i niekompletności programu 4 E (Dąbrowski, 2021). Różne także mogą być kryteria klasyfikacji tych ujęć, na przykład po względem: dynamiki relacji pomiędzy ucieleśnionym podmiotem a otoczeniem, granic systemu poznawczego, aspektów całokształtu poznawczego funkcjonowania, na których koncentruje się dane ujęcie, czy stopnia ucieleśnienia procesów poznawczych.

Jedną z takich klasyfikacji zaproponował Shaun Gallagher (2010). Dokonał jej pod względem stopnia, w jakim – w ramach tych poszczególnych ujęć – ciało wpływa na procesy poznawcze. W ramach swojej klasyfikacji Gallagher wyróżnia pięć takich stopni: 1) ucieleśnienie słabe lub minimalne, które sprowadza się w zasadzie do neuronalnych uwarunkowań procesów poznawczych odseparowanych od środowiska, 2) ucieleśnienie fizjologiczne polegające na wpływie całokształtu budowy organizmu oraz jego możliwości motorycznych na poznanie, 3) ucieleśniona semantyka, która polega na odzwierciedleniu doświadczenia cielesności w strukturach językowych i konceptualizacji świata głównie w formie metafor, 4) ucieleśniony funkcjonalizm, w którym ciało stanowi element szerokiego systemu poznawczego obejmującego mózg, ciało i świat, chociaż można tu mówić o pewnych sytuacjach istnienia niecielesnych reprezentacji wewnętrznych (czysto umysłowych) 5) ucieleśnienie mocne lub radykalne, w przypadku którego całokształt poznania polega na dynamicznych relacjach pomiędzy mózgiem, ciałem i światem.

Współczesne badania dotyczące ucieleśnienia procesów poznawczych obejmują coraz szersze spektrum zagadnień, do których należy na przykład problematyka percepcji, uwagi, języka, pamięci i uczenia się, emocji, a także rozwiązywania problemów, poznania społecznego czy nawet rozumowania moralnościowego. Pojawiają się też różne sposoby operacjonalizacji tych zagadnień, na przykład na gruncie psychologii, co wiąże się z trudnościami w interpretacji gromadzonych wyników badań. Ma to oczywiście także wpływ na sprawę replikacji badań empirycznych w tym zakresie, chociaż kryzys replikacyjny wydaje się w ogóle dotyczyć obszaru nauk o poznaniu. W jakiejś mierze wiąże się to

zapewne z subtelnym charakterem problematyki umysłu jako sfery zdażeń prywatnych.

Z drugiej strony jednakże słyhać właśnie głosy zachęty do wykorzystywania aparatu pojęciowego z zakresu problematyki ucieleśnionego poznania w planowaniu eksperymentów naukowych oraz w interpretacji uzyskiwanych wyników badań (Barrett, 2011). Można odnieść wrażenie, że tak właśnie się dzieje. Rozwój ucieleśnionej kognitywistyki przyczynia się do powrotu na gruncie nauk empirycznych zainteresowania pojęciami, które uważane były przez pewien czas właśnie za zbyt efemeryczne, i wpływa na poszerzenie zakresu podejmowanej problematyki oraz rozwój aparatu pojęciowego w ramach ich systemu wyjaśniającego. Przykład takiego podejścia stanowi chociażby właśnie projekt neurofenomenologii, przedstawiony przez Francisca Varełę (1996/2010) jako propozycja naturalizacji fenomenologii i remedium na kryzys nauk o świadomości. Propozycja ta dotyczy metodologicznego połączenia perspektywy pierwszoosobowej, badanej za pomocą redukcji fenomenologicznej, z metodami trzecioosobowymi – w postaci nowoczesnych technik neuroobrazowania. Efektem takiego podejścia miałyby być wiedza na temat ścisłych i dynamicznych korelacji (*neurophenomenological circulation*) pomiędzy nieznanymi wcześniej elementami struktury doświadczenia a procesami neuronalnymi. Chociaż walor tej propozycji w osiągnięciu swojego głównego celu jako remedium na trudny problem świadomości jest raczej kontrowersyjny, zaproponowany przez Varełę program badawczy doczekał się już konkretyzacji i praktycznych aplikacji, na przykład w badaniach nad neuropsychologicznymi uwarunkowaniami dynamiki ataków wstrząsowych w przypadkach epilepsji i możliwości ich ewentualnej samokontroli (Petitmengin, Navarro & Le Van Quyen, 2007).

Połączenie zatem perspektyw nauk zajmujących się zjawiskami psychicznymi oraz procesami neuronalnymi na gruncie ucieleśnionej kognitywistyki można potraktować jako kolejny etap na drodze do bardziej spójnej i adekwatnej teorii umysłu, który wymaga jednak dalszych wysiłków o charakterze teoretycznym oraz empirycznym.

BIBLIOGRAFIA

- Barrett, L. (2011). *Beyond the Brain: How Body and Environment Shape Animal and Human Mind*. Princeton: Princeton University Press.
- Brooks, R. (1991). Intelligence without Representation. *Artificial Intelligence*, 47(1–3), 139–59. DOI: 10.1016/0004-3702(91)90053-M.
- Clark, A., & Chalmers, D.J. (2008). Umysł rozszerzony. W: M. Miłkowski, & R. Poczobut (Red.), *Analityczna metafizyka umysłu*. Przetł. M. Miłkowski (s. 343–357). Warszawa: Wydawnictwo IFiS PAN.
- Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology Human Perception Performance*. 25(6), 1673–1692. DOI: 10.1037/0096-1523.25.6.1673ht.
- Czerniak, S., & Michalski, R. (2008). *Cielesność, Kompensacja, Mimesis: Wokół pojęciowego instrumentarium współczesnej antropologii filozoficznej*. Warszawa: Wydawnictwo IFiS PAN.
- Dąbrowski, A. (2021), Ucieleśnione poznanie – założenia, tezy i wyzwania. *Argument Biannual Philosophical Journal*, 11(1), 13–32.
- Gallagher, S. (2010), Interpretations of Embodied Cognition. W: W. Tschacher & C. Bergomi (Red.), *The Implications of Embodiment: Cognition and Communication* (s. 59–71). Exeter: Imprint Academic.
- Gibson, J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- Heidegger, M. (2004). *Bycie i czas*. Przetł. B. Baran. Warszawa: PWN.
- Husserl, E. (1975). *Idee czystej fenomenologii i fenomenologicznej filozofii. Księga II*. Przetł. D. Gierulanka. Warszawa: PWN.
- Hutchins, E. (1996). *Cognition in the Wild*. Cambridge: MIT Press.
- Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64, 354–361.
- Maine de Biran, P. (1970). *The influence of habit on the faculty of thinking*. M.D. Boehm (Red.). Westport: Greenwood Press.
- Merleau-Ponty, M. (2001). *Fenomenologia percepcji*. Przetł. M. Kowalska, J. Migasiński. Warszawa: Aletheia.
- Petitmengin, C., Navarro, V., & Le Van Quyen, M. (2007). Anticipating seizure: Pre-reflective experience at the center of neuro-phenomenology. *Consciousness & Cognition*, 16, 746–764.
- Wilson, M. (2008). How did we get from there to here? An evolutionary perspective on embodied cognition. W: P. Calvo, & T. Gomila (Red.), *Handbook of cognitive science: An embodied approach* (s. 375–388). San Diego: Elsevier.

Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*, Cambridge: MIT Press.

Varela, F. (2010). Neurofenomenologia: metodologiczne lekarstwo na trudny problem. Przeł. R. Poczobut, *AVANT*, 1, 31–75.

Ziemke, T. (2015). Czym jest to, co zwiemy ucieleśnieniem? *AVANT*, 3, 161–174. DOI: 10.26913/60202015.0112.0014.

Józef Bremer

Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-9664-8896>

<https://doi.org/10.35765/slowniki.411>

Kognitywistyka społeczna

Streszczenie

DEFINICJA POJĘCIA: W rozdziale przedstawiono definicję kognitywistyki społecznej jako nauki interdyscyplinarnej korzystającej z jednej strony z badań z zakresu filozofii i neuronauk, a z drugiej strony – z interakcjonizmu symbolicznego jako jednego z wiodących nurtów nauk społecznych.

ANALIZA HISTORYCZNA POJĘCIA: Rozdział zawiera opis powstania dwóch obszarów badawczych: 1) kognitywistyki, w skład której wchodzi lingwistyka teoretyczna, psychologia, filozofia, neuronauka, cybernetyka i sztuczna inteligencja, 2) socjologii ze szczególnym akcentem położonym na kształtowanie się interakcjonizmu symbolicznego.

UJĘCIE PROBLEMOWE POJĘCIA: W ciągu ostatnich dziesięcioleci nauki społeczne i kognitywistyka zbliżyły się do siebie. Zarys kognitywistyki społecznej przedstawiono na przykładzie teorii emergencji i neuronauk jako części składowych kognitywistyki. Odwołano się przy tym do uzasadniania wyjaśniającego, traktując kognitywistykę jako naukę, w ramach której tworzy się podstawy do rozwoju kognitywistyki społecznej. W przeciwieństwie do rozwiązań spotykanych w literaturze przedmiotu nie odwołujemy się do mechanicznych wyjaśnień w kognitywistyce społecznej. Teoretyczne ramy tej ostatniej dostarczają zrozumienia funkcjonowania indywidualnego umysłu i zachowania osoby poszerzonego o analizy integracyjne obejmujące poziomy organizacji począwszy od genów, a skończywszy na relacjach społecznych i zjawiskach kulturowych.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Artykuł jest eksploracją kluczowych zagadnień związanych ze społecznymi

aspektami poznania, co rozszerza temat z neuronauki społecznej dotyczącej bezpośrednich interakcji międzyludzkich na cały zakres miejsc, w których zróżnicowanie społeczne oddziałuje z poznawczym. Nacisk położony jest na problemy koncepcyjno-empiryczne wynikające z prób poważnego podejścia do tych zagadnień, a także na nowe zasoby i rozważania związane z takim podejściem. Chodzi jednak także o potrzebę rewizji społecznych koncepcji teoretycznych, aby wykorzystać te zasoby. Artykuł wskazuje na niektóre wnioski – zwłaszcza dotyczące tego, jak proces tzw. socjalizacji należy rozumieć w kategoriach przyjaznych kognitywistyce – a także na problemy natury etycznej związane z jej rozwojem.

Słowa kluczowe: interakcjonizm symboliczny, psychologia społeczna, teorie emergencji, neuronauka społeczna

Definicja pojęcia

Kognitywistyka jest interdyscyplinarną nauką o umyśle, łączącą w swoich badaniach filozofię, neuronaukę, psychologię, lingwistykę teoretyczną i naukę o sztucznej inteligencji. Jej celem jest zrozumienie naszych rozumowań oraz zachowań, a następnie wykorzystanie tej wiedzy chociażby w medycynie, w programach edukacyjnych, w robotyce, w budowaniu coraz bardziej inteligentnych maszyn funkcjonujących na zasadzie symulacji naszej naturalnej inteligencji. Początki tak rozumianej kognitywistyki sięgają połowy lat 50. XX wieku, kiedy badacze z kilku dziedzin nauki połączyli swoje wysiłki w opracowywaniu teorii umysłu opartych na złożonych reprezentacjach mentalnych i procedurach obliczeniowych.

Nauki społeczne skupiają się na zachowaniach jednostek i całych społeczności, na ich wzajemnych interakcjach oraz ich relacjach ze środowiskiem naturalnym i technologicznym. Ich celem jest zrozumienie ewoluujących procesów, w jakie jesteśmy zaangażowani, chociażby w proces zrównoważonego rozwoju naszej planety. Mówiąc o ewolucji współczesnych struktur społecznych, mamy na myśli ich obecność w erze cyfrowej, w społeczeństwie informacyjnym.

Na skrzyżowaniu tych dwóch rodzajów badań powstały i rozwijają się nowe kierunki i programy, takie jak chociażby socjologia poznawcza, psychologia polityczna, ekonomia behawioralna, antropologia poznawcza, kognitywistyka społeczna. Według Tuukka Kaidesoja i jego współpracowników:

Kognitywistykę społeczną można zdefiniować jako dyscyplinę i program badawczy, którego celem jest integracja nauk społecznych (np. socjologii, antropologii, politologii i ekonomii) z elementami kognitywistyki (np. psychologią poznawczą, neuronauką poznawczą itp.) (Kaidesoja, Sarkia & Hyryläinen, 2019, s. 1).

Poniżej po krótkim zarysowaniu historycznego procesu rozwoju kognitywistyki i nauk społecznych przedstawimy bliżej spojrzenie na interakcjonizm symboliczny jako jedną z obecnie wiodących teorii socjologicznych. Z jednej strony da się ją wykorzystać w poszerzaniu zakresu interdyscyplinarności badań kognitywistycznych, a z drugiej strony wyniki tychże badań mogą być użyte w prowadzeniu badań socjologicznych. Odnosząc się do interakcjonizmu, omówimy dwa powiązane ze

sobą projekty kognitywistyki społecznej: jeden z zakresu emergencyjnie rozumianej filozofii umysłu, a drugi z zakresu neuronauk społecznych.

Analiza historyczna pojęcia

Obserwowany współcześnie rozwój kognitywistyki jest ściśle powiązany z tzw. zwrotem poznawczym w badaniach nad umysłem (1950–1960). Do tego okresu główne miejsce w psychologii i filozofii umysłu zajmował behawioryzm, będący reakcją na kwestionowaną wartość badawczą introspekcji stosowanej jako preferowana metoda badań w psychologii. Przyjmowano, że oparte na introspekcji raporty, prowadzone w narracji pierwszoosobowej, na temat wewnętrznego życia psychicznego danej osoby wymykają się zewnętrznej, intersubiektywnej weryfikacji.

Kamieniem milowym w zwrocie kognitywnym było sympozjum na temat teorii informacji, które odbyło się w dniach 10–11 września 1956 roku w Massachusetts Institute of Technology (Cambridge, USA). Spotkali się na nim pionierzy badań sztucznej inteligencji: Allen Newell (1927–1992), Herbert Simon (1916–2001) i Marvin Minsky (1927–2016) oraz językoznawca Noam Chomsky (1928–). Podstawowym założeniem sztucznej inteligencji jest wykorzystanie algorytmów potrafiących symulować ludzkie operacje myślowe i zachowania oraz wykonywać zadania przy użyciu inteligencji podobnej do ludzkiej. Dzięki ludzkiej inteligencji poznajemy, planujemy, przystosowujemy się do nowego otoczenia społecznego, a także potrafimy na nie wpływać.

Natywista Chomsky przeprowadził ostrą krytykę behawiorystycznej koncepcji języka i przedstawił swoją teorię gramatyki uniwersalnej i transformacyjnej. Allen Newell, Herbert Simon i Cliff Shaw zaprezentowali pierwszy program komputerowy służący do prowadzenia automatycznych rozumowań, który nazwano „pierwszym programem sztucznej inteligencji”. Zaprezentowane wieloaspektowe wykorzystanie informatyki zainspirowało neurofizjologa i cybernetyka Warrena McCullocha (1898–1969) oraz logika Waltera Pittsa (1923–1969) do prowadzenia pogłębionych badań nad cyfrowym modelem neuronów. Tym samym opisano funkcjonowanie pierwszej sieci neuronowej. Ponadto istotnymi elementami ożywienia tego rodzaju interdyscyplinarnych badań okazały się sformułowane przez Norberta Wienera (1894–1964) założenia

cybernetyki oraz zaprojektowany przez Alana Turinga (1912–1954) model matematyczny – tzw. maszyna Turinga – służący do realizowania algorytmów, co z kolei przyczyniło się do powstania obliczeniowego modelu umysłu i modularnego modelu mózgu. Z biegiem czasu używanie sztucznej inteligencji odmieniło wiele aspektów naszego życia społecznego: od pojawienia się telefonii komórkowej po opracowanie modeli językowych opartych na architekturze GPT (*Generative Pre-trained Transformer*), czyli kolejne wersje ChatGPT.

Zarysowany na wspomnianym sympozjum program badań był ściśle związany z pracami nad rozwojem sztucznej inteligencji. Jego autorzy prawie bezkrytycznie przejmowali z filozofii akademickiej dominującą wówczas teorię dotyczącą funkcjonowania ludzkiego myślenia (np. teorię identyczności stanów mentalnych i fizycznych, teorię funkcjonalistyczną), traktując je jako narzędzia umożliwiające zamodelowanie stanów intencjonalnych ludzkiego umysłu. Nie zauważali przy tym, że dominujące wśród nich poglądy na temat ludzkiej psychiki mogą być ograniczone w swoim zasięgu (Bremer & Flasiński, 2022). Współczesna kognitywistyka, wsparta nieinwazyjnymi neuroobrazowaniami mózgu, wraca do pytań o świadome przeżycia pierwszoosobowe, o wpływ społeczeństwa i kultury na systemy nerwowe. Wymaga to bardziej skomplikowanych programów badawczych prowadzonych po części w ramach kognitywistyki społecznej.

W większości podręczników przyjmuje się Maxa Webera (1864–1920), Émile’a Durkheima (1858–1917) i Karola Marksa (1818–1883) za twórców socjologii jako odrębnej dyscypliny naukowej. Niemniej istotny wpływ na rozwój nauk społecznych wywarli Charles A. Ellwood (1873–1946) i George Herbert Mead (1863–1931). W swoich pracach inspirowali się filozofią społecznego pragmatyzmu. W roku 1909 Ellwood pisał:

Prawda jest taka, że socjologia i wszystkie inne nauki społeczne muszą poczekać na rozwój biologii i psychologii; a te dawne nauki nawet dzisiaj pozostają w nierozstrzygniętym/nieustalonym stanie. Jakże więc absurdalne jest żądanie od socjologii ustalonego korpusu teoretycznego lub odmowa jej miana nauki! Nie posiada ich żadna z nauk społecznych (Ellwood, 1909, s. 108).

Obydwaj badacze wnieśli wkład w rozwój myśli socjologicznej Herberta Blumera (1900–1987), a tym samym w rozwijany przez niego paradygmat badań socjologicznych, zwany „symbolicznym interakcjonizmem”

(Blumer, 2008), stanowiący określoną perspektywę teoretyczną, umożliwiającą badanie życia grupowego i ludzkich zachowań. Kluczowym mechanizmem kształtowania się struktur społecznych jest bowiem stała wymiana i ewolucja znaczeń, symboli. Zachodzi ona w trakcie wszelkich oddziaływań społecznych między podmiotami, które uświadomiły sobie rolę bycia partnerami. Paradygmat ten jest często traktowany jako teoria i praktyka badawcza zarówno w obszarze psychologii społecznej, jak i socjologii. Za jednego z kolejnych prekursorów symbolicznego interakcjonizmu uznaje się polsko-amerykańskiego socjologa Florianą Znanieckiego (1882–1958), który wraz z Williamem Thomasem (1863–1947) wydał monumentalne dzieło *Chłop polski w Europie i w Ameryce* (Thomas & Znaniecki, 1976).

W latach 90. XX wieku rozwijają się jakościowe metody badań, przyczyniające się do głębszego zrozumienia interakcjonizmu w różnych kontekstach społecznych. Z początkiem XXI wieku teoria ta nadal ewoluje, a w jej ramach badane są nowe interpretacje i zastosowania integrujące symboliczny interakcjonizm ze współczesnymi zagadnieniami socjologicznymi, co pozwala wykazać ciągłe jego znaczenie i możliwości adaptacji (por. Carter & Fuller, 2016). Istotnym obszarem badań podejmowanych w ramach interakcjonizmu symbolicznego jest socjalizacja rozumiana jako procesy, w wyniku których kształtują się osobowości i wyobrażenia o sobie, przekazywane są wartości i postawy, a kultura jednego pokolenia jest powierzana następnemu.

Od strony teoretycznej i praktycznej badania prowadzone w ramach interakcjonizmu symbolicznego mają charakter interdyscyplinarny. W jego obrębie zarówno psychologowie, jak i socjologowie używają terminu „psychologia społeczna”, wskazując na pola specjalizacji swoich dyscyplin. Różnice subdyscyplinarne są widoczne na polu pojęciowym: socjologiczna psychologia społeczna czerpie swoją perspektywę z teorii takich jak symboliczny interakcjonizm Blumera i podkreśla wpływ zmieniających na poziomie indywidualnym na makroskopowe analizy społeczne w badaniu takich tematów jak struktura społeczna i dynamika grupy. Psychologia społeczna uwypukla wpływ kontekstu społecznego na zmienne na poziomie indywidualnym, badając takie efekty jak wpływ społeczny, postrzeganie grupowe i zmiany indywidualnej postawy. W psychologii społecznej podstawy teoretyczne tych i innych efektów zmieniały się często w ciągu ostatnich dziesięcioleci, kładąc nacisk na podejścia

funkcjonalistyczne, behawiorystyczne, atrybucyjne, społeczno-poznawcze, a obecnie multimetodologiczne i politeoretyczne. Różnice metodologiczne są też obecne pomiędzy subdyscyplinami: psychologowie społeczni o nastawieniu socjologicznym chętniej dzisiaj korzystają zarówno z danych socjometrycznych, wywiadów nieustrukturyzowanych, badań ankietowych, technik obserwacyjnych i archiwalnych, jak i z jakościowych metod badawczych opartych na analizie danych zebranych w trakcie wywiadów. Psychologowie społeczni o nastawieniu psychologicznym częściej w swoich dociekaniach opierają się na metodologiach eksperymentalnych. Jednak pomimo różnic w podstawach teoretycznych i metodologii psychologowie społeczni obu tradycji mają w dużej mierze te same zainteresowania badawcze – są nimi ludzkie zachowania indywidualne i społeczne (Uldall, 2013). Jak zobaczymy poniżej, do obydwu z nich odwołują się neuronaukowcy społeczni, wykorzystujący obserwowalne procesy neuronalne (fizjologiczne) i reakcje behawioralne w celu wyciągania wniosków na temat nieobserwowalnych stanów psychicznych. Zakłada się przy tym, że wszystkie poziomy analizy są ze sobą skorelowane (np. stany układu nerwowego wpływają przyczynowo na stany psychiczne, a te mogą w jakiś sposób oddziaływać na reakcje zachodzące na poziomie neuronów). Od swojego powstania interakcjonizm symboliczny otwierał empiryczną perspektywę nauk społecznych dotyczącą badania życia grupowego ludzi i ludzkiego postępowania (Blumer, 2008).

Wśród polskich socjologów prowadzących badania w ramach paradygmatu symbolizmu interakcyjnego należy wymienić przede wszystkim Elżbietę Hałas, spod pióra której wyszło między innymi znaczące opracowanie *Interakcjonizm symboliczny* (2012). W tym opracowaniu, podobnie jak we wcześniejszych oraz w licznych zagranicznych artykułach naukowych, autorka systematyzuje kwestie związane z procesami symbolizacji w zjawiskach życia społecznego, skupiając się między innymi na przedstawieniu związków między kulturą i społeczeństwem. W rozbudowany sposób rozwój interakcjonizmu symbolicznego w Polsce przedstawiają Krzysztof Konecki i Anna Kacperczyk (Konecki & Kacperczyk, 2020). W dalszych rozważaniach będziemy się odwoływać do tej teorii jako jednego z przykładów szerokiej gamy nauk społecznych.

Z końcem wieku XX i na początku wieku XXI nastąpił ogromny rozwój nauk biologicznych, psycholingwistycznych, neuropsychologicznych,

na który oczekiwał Ellwood. Było to możliwe dzięki odkryciu struktury systemu neuronowego (za co w 1906 roku Santiago Ramón y Cajal wraz z Camillem Golgim otrzymał Nagrodę Nobla w dziedzinie medycyny), powstaniu nowych technologii neuroobrazowania, a także szybkim rozwojowi sztucznych sieci neuronowych, zdolnych do uczenia się na podstawie przykładów i potrafiących automatycznie uogólnić nabytą wiedzę. Tym samym nauki społeczne jeszcze ściślej mogły wejść w zakres interdyscyplinarnych badań kognitywistycznych. Rozpoznawanie interakcji pomiędzy agentami, a właściwie zdolność rozpoznawania interakcji społecznych, jest zwykle grupowana z aspektami poznania społecznego wyższego poziomu, takimi jak mentalizacja, czyli zdolność rozumienia zachowań własnych i innych na podstawie stanów psychicznych i skorelowanych z nimi stanów mózgu.

Ujęcie problemowe pojęcia

W cytowanym powyżej artykule Kaideso i jego współpracowników autorzy zaznaczają, że ich celem jest zrekonstruowanie argumentów na rzecz kognitywnych nauk społecznych, określenie ich przesłanek, struktur wnioskowania i wyników oraz ocena tych argumentów według zasadności ich założeń, wiarygodności przesłanek, zasadności wniosków i potencjalnych niejasności pojęciowych. Argumentami tymi są: (1) argument z uzasadnienia wyjaśniającego (*argument from explanatory grounding*), (2) argument z unifikacji teoretycznej (*argument from theoretical unification*), (3) argument z ograniczeń (*argument from constraints*), (4) argument z komplementarności (*argument from complementarity*). Wszystkie cztery argumenty mają uzasadnić potrzebę istnienia kognitywistyki społecznej dla socjologów, chociaż – w odróżnieniu od nas – wspomniani autorzy nie omawiają wyników różnych programów badawczych prowadzonych przez kognitywistów zajmujących się naukami społecznymi (Kaidesoja, Sarkia & Hyyryläinen, 2019). Konkludując, stwierdzają, że udało się im uporządkować argumenty w takiej kolejności, że argument z uzasadnienia wyjaśniającego przemawia za najsilniejszą formą interdyscyplinarnej integracji nauk kognitywnych i społecznych, natomiast argumentacja z komplementarności jest najsłabszą formą uzasadnienia.

Z tej racji poniżej omawiamy szerzej argument z uzasadniania wyjaśniającego, skupiając się – inaczej aniżeli Sarkia wraz ze współpracownikami (Sarkia, Kaidesoja & Hyyryläinen, 2020) – na interakcjonizmie symbolicznym i na jego kognitywistycznym poszerzeniu w ramach nieredukcyjnych teorii emergencji i neuronauki społecznej. Emergencję rozumiemy jako wyłanianie się jakościowo nowych form i zachowań społecznych, będących wynikiem oddziaływania między pojedynczymi osobami z ich umysłami i procesami mózgowymi. Neuronauka społeczna to złożone skrzyżowanie metod i badań psychologii społecznej oraz neuronauk, w ramach której pytamy, jak neuronalne/fizjologiczne procesy oddziałują na zachowanie społeczne. Zarówno zachowania społeczne, jak i skorelowane z nimi procesy neuronalne możemy obserwować i mierzyć empirycznie.

Emergencja zachowań społecznych

W ostatnich dziesięcioleciach zintensyfikowaniu uległo interdyscyplinarne zachodzenie na siebie zakresów oraz metodologii kognitywistyki i nauk społecznych. Na skrzyżowaniu tych dwóch obszarów nauk powstały i rozwinęły się nowe dyscypliny i programy badawcze. Należą do nich socjologia poznawcza, psychologia polityczna, ekonomia behawioralna i nowe programy badawcze z zakresu antropologii poznawczej.

Idąc po linii argumentu z wyjaśniania, przyjmujemy, że kognitywistyce nakierowanej na problemy społecznej należy się miejsce w polu zainteresowań badań nad interakcjonizmem symbolicznym, ponieważ ukazuje ona jego ontologiczne podstawy, a tym samym pogłębia nasze zrozumienie zjawisk społecznych w ogóle. Oznacza to, że jeśli zakładamy, iż od strony wyjaśniania interakcjonizm symboliczny jest zakorzeniony w kognitywistyce, to ta nie może się opierać na zagadnieniach nauk społecznych. Przyjmujemy przy tym znany z filozofii umysłu nieredukcyjny, emergencyjny model wyjaśniający typu: procesy składowe niższego poziomu ontologicznego (fizyczne, chemiczne, biologiczne itp.) jakiegoś złożonego systemu tworzą podstawę dla emergujących (wyłaniających się) z nich procesów wyższego poziomu (ogólnie cech systemowych – mentalnych, społecznych itp.). W teoriach emergencji zazwyczaj przyjmuje się, że cechy wyższego poziomu nie są

sprowadzalne (redukowalne) do obiektów, procesów i cech poziomu niższego – w myśl potocznej zasady: całość (system) jest czymś więcej niż suma jego części. Cechy emergentne danego systemu są ontologicznie nowymi cechami w stosunku do cech podstawowych. O ile przyjmuje się, że obiekty i procesy niższego poziomu oddziałują przyczynowo na procesy, obiekty i cechy wyższego poziomu, o tyle kwestią dyskusyjną pozostaje, czy emergentne cechy wyższego poziomu oddziałują przyczynowo na obiekty niższego poziomu. Chodzi tu o pytanie o tzw. przyczynowość odgórną czy przyczynowość skierowaną ku dołowi, przyczynowość mentalną (*downward causation, mental causation*).

Rozumując analogicznie do biologicznego zjawiska życia, widzianego jako cecha emergentna, i jego relacji do fizyczno-chemicznego podłoża organizmu, możemy w przypadku zjawisk społecznych odróżnić trzy typy przyczynowości skierowanej ku dołowi (Bremer, 2015). Silny typ przyczynowości skierowanej ku dołowi domaga się jasnego rozdzielania poziomu wyższego od podstawowego i ich autonomii. Odpowiada to klasycznemu biologicznemu witalizmowi, przyjmującemu istnienie kreatywnych sił niepodlegających opisowi i wyjaśnianiu mechanicznemu. Przykłady takich teorii można znaleźć w dualistycznie pojmowanej psychologii i filozofii. Niematerialna „siła życia” może przyczynowo wpływać na poziom niższy, czyli na materię, z której składa się ciało danego organizmu.

Zasadniczą cechą umiarkowanego typu omawianej przyczynowości jest, że nie dopuszcza ona, aby cechy wyższego stopnia wywierały bezpośredni, przyczynowy wpływ na obiekty i prawa obowiązujące na stopniu niższym. Oddziaływanie odbywa się poprzez ograniczanie warunków rozwoju na stopniu niższym. Stopień wyższy, skoro już wyemergował, już się wyłonił (lub się wyłania), ogranicza stopień niższy i wpływa na jego funkcjonowanie w odpowiednim zakresie – bez faktycznego, sprawczego fizycznego oddziaływania. Prowadzone przez neurobiologa Jamesa Curleya i jego współpracowników badania nad gryzoniami pokazują, że jakość środowiska społecznego (jako stopień wyższy) może mieć głęboki wpływ na rozwój i aktywność układów nerwowych (stopień niższy), co z kolei ma wpływ na liczne reakcje behawioralne i fizjologiczne, w tym na ekspresję emocjonalności w tymże środowisku (Curley et al., 2011). Autorzy badali zachowania, które ilustrują modulację społeczną podczas rozwoju układów nerwowych, ze

szczególnym uwzględnieniem tych układów, w których obserwuje się efekt długoterminowy. Podobne wyniki odnośnie do wpływu warunków społecznych na systemy nerwowe uzyskano w badaniach nad ssakami naczelnymi, w tym nad ludźmi. U dorosłych mężczyzn, którzy byli maltretowani w dzieciństwie i którzy posiadają gen MAO-A (*Monoamine oxidase A*) o niskiej aktywności, wykazano znacznie większą skłonność do przemocy i zaburzeń w zachowaniach niż u mężczyzn, którzy albo są nosicielami formy wysokiej aktywności owego genu i byli maltretowani, albo którzy są nosicielami w tego genu w postaci o niskiej aktywności, ale nie byli maltretowani (Champagne & Curley, 2005). Wyniki tych badań nie potwierdzają silnego typu przyczynowości, ale przypisujemy im umiarkowany wpływ warunków społecznych na rozwój systemu nerwowego i biologicznego danego osobnika.

Słaby typ biologicznej przyczynowości skierowanej ku dołowi odwołuje się do pojęciowego instrumentarium przestrzeni fazowych dla nieliniarnych, dynamicznych systemów, których trajektorie są przyciągane przez odpowiednie atraktory. Zbudowana na atraktorach teoria poziomów wyższych nie jest wyjaśnieniem tych poziomów, lecz jedynie ich objaśniającym opisem.

W dalszych analizach przyjmujemy tezę o umiarkowanym oddziaływaniu relacji społecznych na nasz system nerwowy. Wracając do zagadnień kognitywistyki społecznej, możemy podsumować, że opieramy się na emergencyjnej, czyli (nieredukcjonistycznej, niemechanicyzycznej wersji) ontologii fizykalistycznej lub inaczej na emergencyjnej ontologii materialistycznej (tzw. słaby materializm). Kluczowym założeniem tej ontologii jest to, że procesy społeczne są zależne od procesów niższego poziomu, ale nie można ich do nich sprowadzić, zredukować, a od strony przyczynowej oddziałują umiarkowanie na procesy w mózgu, z których się wyłoniły (wyemergowały).

Neuronauka społeczna

Poniżej omawiamy przykłady kilku badań z zakresu neuronauki społecznej jako stosunkowo nowej, wyspecjalizowanej części kognitywistyki łączącej symbolizm interakcyjny z naukami o mózgu. Przeprowadzone przez Johna T. Cacioppa i jego współpracowników badania

pokazują, że osobniki społeczne (np. ludzie i pewne gatunki zwierząt) z definicji tworzą wykraczające poza jednostkę struktury, poczynając od tworzenia par i rodzin po większe grupy oraz całe kultury. Struktury społeczne ewoluowały równoległe z mechanizmami neuronalnymi, hormonalnymi, komórkowymi oraz genetycznymi i je wspierały. Pojawiające się dzięki tym struktutom zachowania społeczne pomagały organizmom przetrwać, rozmnażać się, a w przypadku niektórych gatunków opiekować się potomstwem przez wystarczająco długi czas.

Neuronauka społeczna to interdyscyplinarna dziedzina nauki zajmująca się badaniem zagadnień mechanizmów neuronalnych, hormonalnych, komórkowych i genetycznych, a co za tym idzie – badaniem powiązań i wpływów między społecznościami organizmów i biologicznymi poziomami ich organizacji (Cacioppo, Berntson & Decety, 2010). Jako ludzie jesteśmy pod tym względem wyjątkowym gatunkiem społecznym, gdyż nasze instytucje społeczne, cywilizacje i kultury są wysoko rozwinięte, a ich zasięg terytorialny nie zna wąskich granic. Pozostając przy pojęciu emergencji, możemy powiedzieć, że nasza biologia pomogła ukształtować społeczeństwo i środowisko, które stworzyliśmy, a nasze środowisko społeczne pomogło ukształtować nasze ciała i mózgi. Neuronaukę społeczną należy postrzegać jako subdyscyplinę, która ma na celu analizę ludzkich zachowań i biologii oraz zbadanie, w którym miejscu jesteśmy jako gatunek wpisujący się w szerszy łańcuch organizmów biologicznych. Jak wspomniano wcześniej, psychologia społeczna jest także nauką o zachowaniach społecznych, ale w tym wypadku nacisk jest położony na określenie, w jaki sposób myśli, uczucia i zachowanie jednostek wpływają na to, co społecznie rzeczywiste, co wyobrażone lub także na domniemaną obecność innych osób. Percepcja społeczna i poznanie społeczne obejmują poziom intrapersonalny (procesy psychologiczne) oraz interakcje i wpływy społeczne – interpersonalne i grupowe.

Neuronauka społeczna korzysta ze zintegrowanej, wielopoziomowej analizy pojęciowej, która rozciąga się od poziomu neuronowego poprzez podmiotowy do poziomu społecznego. Stwierdza się w niej, że nowe spostrzeżenia i nową wiedzę można uzyskać dzięki uwzględnieniu ustaleń i teorii zaczerpniętych z innych poziomów analiz i badań. W wyniku rozwoju neuronauki społecznej, psychoneuroimmunologii i psychofizjologii coraz bardziej uwydatnia się zdolność wyjaśniająca

integracyjnej analizy. Widać to przy próbach modelowania złożonych zjawisk, które rozwijają się na różnych poziomach organizacji neurologicznej. Wysiłki mające na celu integrację informacji na różnych poziomach analiz są szczególnie trudne, biorąc pod uwagę nieodłączną złożoność systemów neurobiologicznych z jednej strony i symbolizmu interakcyjnego z drugiej strony. Ale tylko dzięki takim badaniom możliwe będzie pełniejsze, kognitywistyczne zrozumienie zarówno zagadnień z zakresu relacji umysł – ciało, jak i zagadnień z zakresu symbolicznego interakcjonizmu.

Michael P. Kelly i jego współpracownicy stawiają hipotezę dotyczącą wpływu interakcji społecznych na modele przetwarzania predykcyjnego w mózgu (Kelly et al., 2019). Według badaczy kodowanie predykcyjne zakłada, że mózg aktywnie przewiduje nadchodzące bodźce zmysłowe, zamiast je jedynie pasywnie rejestrować. Kodowanie predykcyjne jest efektywne w tym sensie, że mózg nie musi utrzymywać wielu wersji tej samej informacji na różnych poziomach hierarchii przetwarzania. Modele te integrują koncepcje postrzegania pochodzące z neuronauki oraz z socjologii, uwidaczniając zarazem rozbieżności w sposobach opisywania zjawiska angażowania się myślących podmiotów w świat społeczny. Autorzy artykułu łączą koncepcje modeli przetwarzania predykcyjnego z Weberowskimi typami idealnymi i praktyką społeczną, w tym z tezami interakcjonizmu symbolicznego. W ten sposób powstaje w miarę ujednoczona hipotetyczna struktura integrująca postrzeganie świata społecznego i hipotetyczne procesy mózgowe. Hipoteza ta łączy aspekty neuronauki i psychologii z teorią społeczną, z zamiarem pokazania, w jaki sposób zachowania społeczne można kognitywistycznie „odwzorować” na procesy mózgowe. Tym samym otrzymujemy ramy koncepcyjne łączące obie dyscypliny i otwierające między nimi twórczy dialog oraz perspektywę przyszłych badań.

Związany z technologiami neuroobrazowania postęp dokonujący się w psychologii społecznej i neuronauce umożliwił badanie różnicowanej aktywacji poszczególnych obszarów mózgu podczas normalnego i zaburzonego myślenia, a tym samym odkrywanie złożonych struktur ludzkiego umysłu. Ponadto interfejsy typu mózg – komputer osiągnęły dzisiaj taki poziom zaawansowania, dzięki któremu chirurgicznie wszczepione implanty domózgowe (implanty neuronowe, neuroprotezy) są w stanie „przełożyć” wywołaną myślami aktywację mózgu na ruch

w urządzeniach zewnętrznych. Elektrody umieszczane w mózgu w celu rejestrowania lub stymulowania aktywności neuronowej otwierają nowy, ekscytujący obszar badań w zakresie neuronauki, inżynierii biomedycznej i nauk społecznych. Tego typu innowacyjne techniki rewolucjonizują nasze rozumienie mózgu i wskazują na nowe możliwości usprawniania funkcjonowania neuronów. Dzięki bezpośredniemu połączeniu z obwodami nerwowymi mózgu implanty domózgowe oferują różne zastosowania, w tym przywracanie osobie funkcji sensorycznych i zwiększanie zdolności poznawczych. Anubhav Singh omawia postęp w dziedzinie implantów mózgowych i ich potencjalny wpływ na ludzkie zachowania i samopoczucie, a tym samym na interakcje społeczne. Innowacje te są niezwykle obiecujące dla osób z zaburzeniami wzroku, mowy lub paraliżem, ponieważ oferują im możliwość skutecznej i niezależnej komunikacji społecznej (Singh, 2023).

Te i inne osiągnięcia w badaniach kognitywistycznych reprezentują nową erę w rozumieniu biologicznych i psychologicznych podstaw naszej ludzkiej egzystencji. Korzystając z analogii z zakresu fizyki, możemy powiedzieć, że obserwacje dualizmu korpuskularno-falowego jako cechy obiektów kwantowych (np. fotonów czy elektronów), polegającego na przejawianiu w zależności od sytuacji właściwości falowych (dyfrakcja, interferencja) lub korpuskularnych (określona lokalizacja, pęd) wykazały nieadekwatność fizyki klasycznej i zapoczątkowały powstanie fizyki kwantowej. Podobnie coraz bardziej oczywiste staje się, że konwencjonalne, monodyscyplinarne modele zarówno w biologii, jak i psychologii społecznej okazują się nieodpowiednie lub przynajmniej niekompletne przy próbach wyjaśniania złożoności zachowań indywidualnych i społecznych.

Od kilku lat Polska Akademia Nauk we współpracy z innymi podmiotami prowadzi w Warszawie Pracownię Neuronauki Społecznej, której głównym celem jest badanie procesów neuronalnych i behawioralnych, kształtujących sposoby, w jakie przetwarzamy sygnały społeczne (<https://psych.pan.pl/struktura/pracownie-naukowe-ip-pan/pracownia-neuronauki-spoecznej/>). Prowadzone obecnie badania empiryczne dotyczą pytania o to, jakie czynniki wpływają na nasze społeczne procesy poznawcze i leżącą u ich podstaw społeczną aktywność mózgu. Badane są społeczne zdolności poznawcze i aktywność „mózgu społecznego” zarówno u osób z populacji ogólnej, jak i u pacjentów

z zaburzeniami neuropsychiatrycznymi, aby uchwycić czynniki wpływające na społeczne funkcjonowanie człowieka. Badania te skupiają się zwłaszcza na wpływie poczucia samotności (lub technicznie: „postrzeganej izolacji społecznej”) na procesy behawioralne i fizjologiczne związane z poznaniem społecznym.

Z jednej strony badania neuronauk społecznych, w tym implanty mózgowe, oferują liczne możliwości rozwoju medycyny i poprawy jakości życia ludzi, a tym samym przywracają społeczeństwu zdrowych lub zdrowszych obywateli. Z drugiej strony, co niestety rzadko jest podnoszone w dyskusjach lub opracowaniach naukowych, ich stosowanie budzi istotne obawy natury etycznej. Chodzi tu zwłaszcza o kwestie związane z zachowaniem prywatności i bezpieczeństwa badanych osób. W przypadku interfejsów typu mózg – komputer istnieje możliwość nieautoryzowanego, nieuświadomionego ze strony pacjentów dostępu do ich danych neuronowych. Rodzi to pytania o ochronę tak pozyskanych danych i zgodę na ich wykorzystanie. Istotne staje się znalezienie równowagi między korzyściami płynącymi z udostępniania informacji dla postępu naukowego a ochroną prywatności. Co więcej, zwiększanie ludzkich zdolności i możliwości poznawczych za pomocą implantów wszczepianych do mózgu wywołuje dyskusje dotyczące kwestii sprawiedliwości i równości społecznej. Jeśli neuroprotezy staną się łatwo dostępne, istnieje ryzyko poszerzania się przepaści społecznej między tymi, którzy mogą sobie na to pozwolić lub mają dostęp do takich ulepszeń, a tymi, którzy takich możliwości nie mają. Zapewnienie społecznie sprawiedliwej dystrybucji i dostępności finansowej tego rodzaju technologii będzie miało zasadnicze znaczenie dla zapobiegania pogłębianiu się już istniejących różnic społecznych.

Refleksja systematyczna z wnioskami i rekomendacjami

Z jednej strony od lat 70. minionego wieku w kognitywistyce nastąpił zwrot społeczny, na co wskazuje pojawienie się neuronauki społecznej i innych eksperymentalnych programów badawczych, z drugiej strony wielu badaczy społecznych o orientacji teoretycznej i empirycznej w większym stopniu zainteresowało się kognitywistyką i czerpało z niej

informacje, na co przykładowo wskazuje powstanie socjologii poznawczej czy psychologii społecznej. W artykule przedstawiono i rozwinięto kognitywistyczne podejście do interakcjonizmu symbolicznego jako nauki społecznej. Uzyskane podstawowe pojęcia i spostrzeżenia teoretyczne dotyczące interakcjonizmu okazują się istotne dla szeroko rozumianej socjologii oraz psychologii społecznej. Tradycyjne dyscypliny nauk o organizmach żywych zajmują się złożonymi zjawiskami zachodzącymi na danym im poziomie opisu i wyjaśniania, czy to na poziomie molekularnym, anatomicznym organizmu, czy na poziomie psychologicznym lub społecznym. Uzyskane w ramach ich badań wyniki dostarczają kluczowych informacji na temat procesów biologicznych, psychologicznych oraz społecznych. Kognitywistyczne, interdyscyplinarne podejście cechuje się wielopoziomowością analiz pojęciowych połączonych z badaniami empirycznymi i niesie ze sobą dodatkową wartość wyjaśniającą.

Zachowania społeczne można zrozumieć na bazie filozoficznej teorii emergencji z umiarkowaną wersją przyczynowości skierowanej ku dołowi. Podobnie jak teoria ta zapobiega redukcji stanów mentalnych do stanów fizycznych (mechanicznych), tak samo jej społeczna wersja zapobiega redukcji zachowań społecznych do stanów fizycznych. Teorie emergencji możemy traktować jak surogaty teorii neuronaukowych. Emergentystyczne ujęcie stanów społecznych odpowiada tezom z neuronauki społecznej, interpretacja której również nie jest skazana na redukcję stanów społecznych do stanów neuronalnych. Neuronauka społeczna wskazuje na metodologiczną konieczność dostrzegania istotności związku między wydarzeniami zachodzącymi na poziomie neurobiologicznym a zachowaniami społecznymi (opisywanymi w języku nauk społeczno-humanistycznych używającym kategorii językowych trzecio- i pierwszoosobowych). Te ostatnie mogą się opierać na wydarzeniach opisywanych w języku interakcyjnego symbolizmu. Opisy te skłoniły naukowców zajmujących się biologią, kognitywizmem i naukami społecznymi do współpracy w bardziej systematyczny sposób. Towarzystwo temu powszechne przekonanie, że zrozumienie funkcjonowania indywidualnego umysłu i zachowania osoby można poszerzyć za pomocą analizy integracyjnej obejmującej poziomy organizacji począwszy od genów, a skończywszy na relacjach społecznych i zjawiskach kulturowych. W ostatnich latach obserwujemy ekspansywny rozwój

neuronauk, które obejmują społeczną neuronaukę poznawczą, społeczną neuronaukę afektywną i kulturową, społeczną neuronaukę obliczeniową, rozwojową i porównawczą.

Neuronauka społeczna wykorzystuje na sposób metodologiczno-teoretyczny dane z badań mózgu do opisywania i wyjaśniania procesów społecznych. Ogólne podejście neuronauki społecznej, jako części kognitywistyki, zmierza do zrozumienia procesów psychologicznych leżących u podstaw naszych zachowań społecznych. Ponieważ zachodzące w danej osobie procesy psychiczne są zjawiskami, których nie można bezpośrednio zaobserwować, neuronaukowcy społeczni opierają się z jednej strony na kombinacji mierzalnych lub obserwowalnych reakcji neuronalnych i fizjologicznych, a z drugiej strony równocześnie na obserwacji naszych potocznych zachowań. Stąd wyciągają wnioski na temat stanów psychicznych. Badając reakcje neuronalne, neuronaukowcy społeczni mogą w przyszłości powiedzieć nam więcej nie tylko o zachowaniach społecznych czy o obszarach mózgu skorelowanych z tymi zachowaniami, ale także o tym, w jaki sposób procesy społeczne wpływają na funkcjonowanie systemów nerwowych.

Wspomniane badania nad implantami mózgowymi niosą za sobą możliwość adekwatniejszego zrozumienia mózgu, zrewolucjonizowania metod leczenia, a tym samym przywracania osobom sprawności fizycznej i umysłowej, co z kolei umożliwi ich społeczną integrację. Od przywracania utraconej funkcjonalności po zwiększanie zdolności poznawczych, interfejsy neuronowe oferują nowe możliwości ulepszenia i interakcji między ludźmi. Pojawiają się jednak pytania: Jakie są lub będą społeczne i kulturowe konsekwencje wykorzystania tego rodzaju technologii umożliwiających manipulowanie naszymi mózgami, a co za tym idzie, pośrednio naszymi umysłami i zachowaniami społecznymi? Jaki wpływ neuronauka społeczna będzie miała na nasze samorozumienie i na ogólne, społeczne pojęcie o człowieku? W ślad za implementowaniem nowych technologii istotne jest uwzględnianie kwestii etycznych związanych z poszanowaniem prywatności, równouprawnienia i dostępności tego rodzaju narzędzi. W miarę postępu niezbędne jest wspieranie interdyscyplinarnej współpracy między naukowcami, etykami, decydentami i przedstawicielami grup społecznych, aby zapewnić opracowywanie i wdrażanie implantów mózgowych w sposób odpowiedzialny, etyczny i korzystny dla całego społeczeństwa. Utrzymując

właściwą równowagę, możemy uwolnić niezwykley potencjał, jaki niosą ze sobą implanty mózgowy, chroniąc jednocześnie podstawowe wartości i dobro jednostek i społeczeństwa jako całości.

BIBLIOGRAFIA

- Blumer, H. (2008). *Interakcjonizm symboliczny. Perspektywa i metoda*. Przeł. G. Woroniecka. Kraków: Zakład Wydawniczy Nomos.
- Bremer, J. (2015). Przyczynowość skierowana ku dołowi i jej rozumienie w biologii. *Poznańskie Studia z Filozofii Nauki*, 24(1), 93–115.
- Bremer, J., & Flasiński, M. (2022). The Turing Test, or a Misuse of Language when Ascribing Mental Qualities to Machines. *Forum Philosophicum*, 27(1), 6–25. DOI: 10.35765/forphil.2022.2701.01.
- Cacioppo, J.T., Berntson, G.G., & Decety, J. (2010). Social Neuroscience and Its Relationship to Social Psychology. *Social Cognition*, 28(6), 675–685. DOI: 10.1521/soco.2010.28.6.675.
- Carter, M.J., & Fuller, C. (2016). Symbols, Meaning, and Action: The Past, Present, and Future of Symbolic Interactionism. *Current Sociology*, 64(6), 931–961. DOI: 10.1177/0011392116638396.
- Champagne, F.A., & Curley, J.P. (2005). How Social Experiences Influence the Brain. *Current Opinion in Neurobiology*, 15(6), 704–709. DOI: 10.1016/j.conb.2005.10.001.
- Curley, J.P., Jensen, C.L., Mashoodh, R., & Champagne, F.A. (2011). Social Influences on Neurobiology and Behavior: Epigenetic Effects during Development. *Psychoneuroendocrinology*, 36(3), 352–371. DOI: 10.1016/j.psyneuen.2010.06.005.
- Ellwood, Ch.A. (1909). The Science of Sociology: A Reply. *American Journal of Sociology*, 15(1), 105–110. Pobrane z: <https://www.jstor.org/stable/2762623> (dostęp: 06.03.2024).
- Hałas, E. (2012). *Interakcjonizm symboliczny. Społeczny kontekst znaczeń*. Warszawa: Wydawnictwo Naukowe PWN.
- Kaidesoja, T., Sarkia, M., & Hyryläinen, M. (2019). Arguments for the Cognitive Social Sciences. *Journal for the Theory of Social Behaviour*, 49(4), 480–498. DOI: 10.1111/jtsb.12226.
- Kelly, M.P., Kriznik, N.M., Kinmonth, A.L., & Fletcher, P.C. (2019). The Brain, Self and Society: A Social-Neuroscience Model of Predictive Processing. *Social Neuroscience*, 14(3), 266–276. DOI: 10.1080/17470919.2018.1471003.
- Konecki, K.T., & Kacperczyk, A. (2020). Symbolic Interactionism in Poland: Inspirations and Development. *Qualitative Sociology Review*, 16(4), 8–34. DOI: 10.18778/1733-8077.16.4.02.

- Patel, H. (2021). Fundamental Principles of Artificial Intelligence. *International Research Journal of Modernization in Engineering Technology and Science*, 3(1), 318–323. Pobrano z: https://www.irjmets.com/uploadedfiles/paper/volume3/issue_1_january_2021/5605/1628083227.pdf (dostęp: 06.03.2024).
- Sarkia, M., Kaidesoja, T., & Hyyryläinen, M. (2020). Mechanistic explanations in the cognitive social sciences: Lessons from three case studies. *Social Science Information*, 59(4), 580–603. DOI: 10.1177/0539018420968742.
- Singh, A. (2023). Exploring the Potential of Brain Implants: Unlocking the Secrets of Neural Enhancement. *Journal of Neuroscience and Neuropharmacology*, 9(1), 001–002. DOI: 10.4172/24699780.2023.9.1.189.
- Thomas, W.I., & Znaniecki, F. (1976), *Chłop polski w Europie i Ameryce*. T. 1–5. Warszawa: Ludowa Spółdzielnia Wydawnicza.
- Uldall, B.R. (2013). Social Psychology. W: A. Runehov, & L. Oviedo (Red.), *Encyclopedia of Sciences and Religions* (s. 2171–2179). Dordrecht: Springer. DOI: 10.1007/978-1-4020-8265-8_1047.

Ewa Jaglarz

Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-8754-8770>

<https://doi.org/10.35765/slowniki.455>

Zaburzenia mowy i języka

Streszczenie

DEFINICJA POJĘCIA: W opracowaniu zaprezentowano różne ujęcia rozumienia zaburzeń mowy i języka poczynawszy od medycznych, poprzez behawioralne, objawowe, aż do kognitywistycznych. W pierwszej części podjęto problematykę rozróżnienia kategorii języka, mowy i mówienia. Język w kontekście psycholingwistycznym i kognitywistycznym ujęto jako system znaków, byt o charakterze mentalnym, stosowany do celów komunikacyjnych i społecznych, mowę zaś jako wieloaspektowy, jednostkowy akt użycia tego systemu.

ANALIZA HISTORYCZNA POJĘCIA: Pierwsze ślady zainteresowania problematyką odnajdujemy w praktyce lekarskiej foniatorów. Później dzięki aktywności powstających ośrodków naukowych nastąpiło wyodrębnienie teorii zaburzeń jako odrębnej dyscypliny nauki. Dziś teoria zaburzeń mowy i języka ma swoje zastosowanie również w kognitywistyce. Problematykę tę podjęło kilka ośrodków naukowych w Polsce.

UJĘCIE PROBLEMOWE POJĘCIA: Złożona natura języka i mowy, heterogeniczność konwergencji i dywergencji objawów i przyczyn oraz interdyscyplinarność problematyki utrudnia lub uniemożliwia dokonanie precyzyjnego rozgraniczenia rodzajów zaburzeń mowy i języka. Dokonano krytycznego przeglądu teoretycznych koncepcji zaburzeń, ujawniając ich różnicowany charakter szczególnie w wymiarze kognitywistycznym.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Zaprezentowana analiza zaburzeń znakomicie odzwierciedla wymagania w zakresie diagnozy i zrozumienia ich różnych przejawów i form. Język

poprzez wielorakość funkcji jest narzędziem niezbędnym dla podmiotowego uczestnictwa osobnika w życiu społecznym. Ujęcie kognitywistyczne pozwala na pragmatyczne zastosowanie wiedzy o zaburzeniach mowy i języka.

Słowa kluczowe: język, mowa, zaburzenia mowy i języka, psycholingwistyka, kognitywistyka

Definicja pojęcia

Język jest złożonym systemem komunikacji, który charakteryzuje ludzką specyfikę i umożliwia interakcje społeczne. Jest to narzędzie umożliwiające ludziom wyrażanie myśli, emocji, przekonań i intencji za pomocą jasno zdefiniowanych struktur syntaktycznych, semantycznych oraz fonologicznych. Język można rozpatrywać z różnych perspektyw – lingwistycznej, psychologicznej, antropologicznej, neurobiologicznej i kognitywistycznej. W kognitywistyce język jest postrzegany przede wszystkim jako narzędzie umysłu, umożliwiające przetwarzanie informacji, myślenie i komunikację. Kognitywiści badają, jak język jest reprezentowany i przetwarzany w mózgu, a także jak jest używany w procesie myślenia. Niektórzy podkreślają znaczenie języka w kształtowaniu procesów myślowych. Hipoteza Sapira-Whorfa sugeruje, że język, którym się posługujemy, wpływa na sposób, w jaki postrzegamy i interpretujemy świat wokół nas. Psycholingwistyka bada procesy psychiczne odpowiedzialne za produkcję języka ale też generowanie i rozumienie mowy oraz pisanie. W podejściu opartym na użyciu (*Usage-Based Approach*) zakłada się, że nabywanie języka i jego struktury są kształtowane przez doświadczenia językowe. W tym ujęciu język jest plastyczny i dynamicznie się rozwija w zależności od kontekstu komunikacyjnego. Neurolingwistyka koncentruje się na biologicznych podstawach przetwarzania języka, w tym na tym, jak różne obszary mózgu przyczyniają się do rozumienia i produkcji języka. Kognitywistyka jako dziedzina interdyscyplinarna pozwala na szerokie rozumienie zaburzeń mowy i języka.

Język, będący jednocześnie produktem i narzędziem ludzkiego umysłu, jest zjawiskiem wielowymiarowym, łączącym w sobie aspekty symboliczne, kulturowe, poznawcze i biologiczne. Badania nad językiem przyczyniają się nie tylko do zrozumienia samego języka, ale także ludzkiego umysłu, społeczeństwa i kultury, dlatego mają multidyscyplinarny i interdyscyplinarny charakter.

Ogólnie rzecz ujmując, zaburzenia języka, rozumiane również jako zaburzenia komunikacji językowej, przyjmują formę dysfunkcji komunikacyjnej, która wpływa na zdolność jednostki do rozumienia, przetwarzania i produkowania języka mówionego i pisanego w jednej lub więcej jego fundamentalnych domenach: fonologii, morfologii, składni, semantyki oraz pragmatyki. Są to zaburzenia, które mogą mieć różne

przyczyny i manifestować się na wiele sposobów w zależności od wieku, doświadczenia językowego i kontekstu kulturowego osoby. Mogą one obejmować zaburzenia nabyte, takie jak afazja, która często jest wynikiem uszkodzenia mózgu, oraz zaburzenia rozwojowe, takie jak specyficzne zaburzenia rozwoju językowego (SLI). Wpływ tych zaburzeń może się różnić, prowadząc do trudności w nauce, interakcjach społecznych i na gruncie zawodowym.

Zaburzenia mowy i zaburzenia języka są powiązane, ale według niektórych poglądów różnią się pod względem obszarów, które wpływają na komunikację. W literaturze przedmiotu rozróżnienie pojęć „język” i „mowa” przypisuje się Ferdinandowi de Saussure, szwajcarskiemu językoznawcy, uważanemu za ojca współczesnej lingwistyki strukturalnej. W swoim podejściu do językoznawstwa, które przedstawił w pracy *Kurs językoznawstwa ogólnego*, de Saussure zdefiniował „język” (*langue*) jako społeczny system znaków i reguł, który jest wspólny dla członków danej społeczności językowej. Natomiast „mowa” (*parole*) to w jego ujęciu indywidualne użycie języka przez człowieka, czyli sposób, w jaki poszczególni ludzie i grupy używają tego systemu w komunikacji. Szwajcarski językoznawca skupiał się przy tym na strukturalnych i semiotycznych aspektach języka, wprowadzając na przykład pojęcie „znaku językowego” oraz rozróżnienie dwóch jego aspektów: formy słowa lub wyrażenia (*signifiant*) oraz znaczenia, które to pojęcie reprezentuje (*signifié*). Zwracał ponadto uwagę na to, że znaczenia w języku wiążą się ściśle z różnicami pomiędzy jego elementami. Słowa zyskują zatem swoje znaczenie nie w izolacji, lecz w ścisłej relacji z innymi słowami.

Strukturalne podejście de Saussure’a uzupełnione zostało o aspekty funkcjonalne przez Romana Jakobsona, który zaproponował kompleksowy model komunikacji, obejmujący na przykład fonologię i uwzględniający różne funkcje języka w różnych kontekstach. Jakobson nie podważał samej istoty rozróżnienia na normy (*langue*) i aktualności mowy jednostkowej (*parole*). Zakładał, że normy językowe są społeczne, a mowa jednostkowa, normy są niezmiennie, a wszystko, co zmienne, pojawia się w mowie jednostkowej. Normy są jednolite, homogeniczne, a mowa heterogeniczna i fizjologiczna. Dla Jakobsona językoznawstwo dzieli się na dwie dyscypliny – jedna zajmuje się normami społecznymi, które stanowią system statyczny, możliwy do opisanego jako wzajemnie

powiązana całość, druga opisuje mowę jednostkową postrzeganą jako przypadkowa i indywidualna (Jakobson, 1983). W odróżnieniu jednak od de Saussure'a, który analizował język takim, jaki jest „tu i teraz”, skupiając się na strukturach istniejących w danym momencie (analiza synchroniczna), Jakobson rozwijał podejście, które łączyło analizę synchroniczną z analizą diachroniczną. Analizował język w konkretnym użyciu, ale również opisywał, jak język zmienia się i rozwija na przestrzeni czasu. Owo przeciwstawienie „wirtualności” norm języka i aktualności mowy jako jednostkowej niewątpliwie dało przyczynek do pogłębionych badań nad zaburzeniami mowy. Zresztą Jakobson sam zajmował się zaburzeniami języka – afazją.

W tym kontekście należy przywołać poglądy Noama Chomsky'ego, zaprezentowane w latach 60. ubiegłego wieku. Amerykański lingwista znany jest głównie z teorii gramatyki uniwersalnej, która zakłada, że wszyscy ludzie rodzą się z wrodzoną zdolnością do nauki języka. Według tej teorii wszystkie języki świata mają wspólną podstawową strukturę, a różnice między nimi są tylko powierzchowne. Ta wrodzona, nieuświadomiona zdolność językowa, nazwana przez autora „kompetencją językową”, pozwala dzieciom szybko i skutecznie przyswajać język ojczysty. Chomsky rozwinął również teorię transformacyjno-generatywnej gramatyki, która opisuje, jak zasady gramatyczne mogą być używane do generowania prawidłowych zdań w języku i powodować, że z ograniczonego zestawu reguł możemy wytworzyć niemal nieskończoną liczbę zdań. Równoległe Chomsky wprowadził pojęcie „performancji” – wykonania, aby rozróżnić wewnętrzną wiedzę językową użytkownika języka (kompetencję) oraz rzeczywiste użycie języka w konkretnych sytuacjach. Performancja obejmuje więc faktyczne wypowiedzi, ich konstrukcję i interpretację w kontekście codziennej komunikacji. Według Chomsky'ego performancja jest często obarczona błędami i ograniczeniami, wynikającymi z czynników psychologicznych, społecznych i środowiskowych, które wpływają na sposób, w jaki ludzie używają języka (Greene, 1977). W odróżnieniu od de Saussure'a, który traktował język jako byt obiektywny, u Chomsky'ego widzimy język przetwarzany i używany, który pojawia się w umyśle dzięki abstrakcyjnym, wrodzonym strukturom i mechanizmom umysłowym – bytom mentalnym.

Ida Kurcz na gruncie polskich badań psycholingwistycznych ujmuje język jako system znaków i reguł nimi zarządzających, a mowę jako

konkretne akty użycia tego systemu (Kurcz, 2000, s. 14). Poza tym autorka zwraca uwagę na odrębność języka od mowy i mówienia. W jej rozumieniu mowa to

zjawisko wzajemnego porozumiewania się ludzi głównie za pomocą dźwięków (wtórnie za pomocą gestów czy też w formie pisanej itp.). Mówienie zaś to konkretne wypowiedzi, konkretne akty użycia języka przez konkretnych ludzi, to właśnie de Saussurowskie „parole”. Język zaś jest systemem znaków i reguł posługiwania się nimi. Jest więc czymś, co jest stosowane, używane w mowie i realizuje się w konkretnych aktach mówienia (Kurcz, 1987, s. 11).

Podobne rozumienie można znaleźć w uwagach Stanisława Grabiasa, choć we własnej definicji akcentuje on funkcjonalny aspekt języka:

mowa jest układem czynności, jakie przy udziale języka wykonuje człowiek, poznając świat i przekazując wiedzę o sobie i o świecie innym uczestnikom życia społecznego. Definicja ta wynika z przekonania, że język wiedzie wprost do umysłu człowieka: jego możliwości poznawczych, emocji i chcień oraz że zachowania językowe pozwalają też oceniać aktywność ludzką poprzez opis możliwości przekazu wiedzy i sposobu istnienia jednostki w grupie społecznej (Grabias, 2020, s. 534).

W ramach zaprezentowanego tu metateoretycznego ujęcia mowa i język stanowią problematykę wieloaspektową, prezentowaną przez poszczególnych badawczy za pomocą różnego rozkładu akcentów. W podejściach tych zwraca się uwagę na strukturalny i funkcjonalny aspekt tej problematyki, różne elementy tej struktury i jej różne funkcje, aspekty niezmiennicze oraz dynamiczne, ich uwarunkowania (społeczne, biologiczne, poznawcze) itd.

Ta różnorodność zasadniczych ujęć problematyki mowy i mówienia przekłada się na zróżnicowanie podejść do kwestii ich zaburzeń. Z jednej strony możemy uznać, że zaburzenia mowy dotyczą trudności z fizycznym procesem mówienia, fizyczną, biologicznie uwarunkowaną realizacją języka. Mogą one obejmować problemy z artykulacją (czyli precyzyjnym tworzeniem dźwięków), płynnością mowy (jąkanie) oraz modulacją głosu, w tym tonem, głośnością i jakością głosu. Zaburzenia mowy skupiają się na mechanicznym aspekcie wytwarzania mowy, choć podłoże zaburzeń może być bardzo zróżnicowane. Mogą mieć przyczyny takie jak: uszkodzenia neurologiczne, wady rozwojowe, urazy, choroby oraz czynniki genetyczne. Diagnoza i leczenie tych zaburzeń

często wymagają interdyscyplinarnego podejścia, włączając specjalistów – logopedów, neurologów, psychologów i psycholingwistów. Zaburzenia mowy można uznać za grupę zróżnicowanych stanów patologicznych dotyczących zdolności produkowania mowy, co wpływa na sposób, w jaki jednostka wyraża siebie za pomocą języka werbalnego.

Z drugiej strony zaburzenia języka odnoszą się do trudności z rozumieniem lub używaniem słów w kontekście i obejmują zarówno język mówiony, jak i pismo. Zaburzenia te mogą wpływać na zdolność do formułowania zrozumiałych zdań, używania odpowiedniego słownictwa, rozumienia poleceń czy też budowania złożonych struktur językowych. Zaburzenia językowe mogą dotyczyć ekspresji (zdolności do wyrażania myśli i idei) oraz rozumienia (zdolności do rozumienia mowy i tekstu).

W paradygmacie zaburzeń Stanisław Grabias dokonał swoistej syntezy dwóch pojęć, jakimi są język i mowa. Wykorzystując dedukcyjne pośrednie wnioskowanie, uznał, że przejawem występowania zaburzeń mowy będzie obserwowana trudność z nabyciem języka:

Pojęcie zaburzeń mowy winno wynikać z analizy takich sytuacji, w których człowiek nie jest w stanie opanować języka w ogóle lub opanowuje go tylko trochę bądź takich sytuacji, w których znając język nie może budować żadnych wypowiedzi językowych lub zmuszony jest budować wypowiedzi zaburzające interakcję społeczną (Grabias, 2012, s. 24).

Dostrzec tu można wyraźnie wzajemny związek przytoczonych kategorii. Autor posunął się jeszcze dalej, utożsamiając mowę z komunikacją językową, tak by móc postrzegać zaburzenia mowy na tle pełnego obrazu komunikacji językowej „pozwole bowiem dostrzegać nie tylko to, co w związku z określoną dysfunkcją istnieje (...), ale również i to, czego nie ma, a co powinno istnieć, aby zachowania językowe były skuteczne” (Grabias, 2012, s. 49).

Jak widać, obie kategorie – języki i mowa – są ważne dla komunikacji i mogą się wzajemnie przenikać. Na przykład dziecko z zaburzeniami mowy może mieć trudności z wyraźnym mówieniem, co z kolei może wpływać na jego zdolność do komunikowania się i nauki języka. Podobnie problemy z rozumieniem języka mogą utrudniać naukę poprawnej artykulacji i innych aspektów mowy, które posiadają istotny status w aktywności komunikacyjnej ujmowanej również w kategoriach procesu poznawczego.

Analiza historyczna pojęcia

W historii badań nad zaburzeniami mowy już od starożytności upatrywano związki między mózgiem a funkcjami mowy (przykład rozważań Hipokratesa i Arystotelesa). Prekursorem nowożytnych badań na ten temat był francuski neurolog Paul Broca, wskazujący na anatomię funkcjonalną mózgu jako neuronalne podłoże mowy, które w przypadku uszkodzeń warunkuje jej zaburzenia. Podobnie niemiecki neurolog Carl Wernicke uważany jest za odkrywcę mózgowej lokalizacji funkcji rozumienia mowy, a Adolf Kussmaul wprowadza termin „afazja” i klasyfikuje różne typy tego zaburzenia. Prace nad mapowaniem mózgu kontynuowane były w XX wieku przez Wildera Penfielda. XX wiek to również intensywny rozwój psycholingwistyki i neuropsychologii. Noam Chomsky wprowadza pojęcie gramatyki generatywnej, sugerując, że zdolność do nauki języka jest wrodzona, Ferdinand de Saussure przez dostarczenie narzędzi do analizy języka jako systemu wywiera istotny wpływ na rozwój psycholingwistyki, a prace na temat języka i fonologii Romana Jakobsona przyczyniają się do rozwoju badań nad zaburzeniami mowy w zakresie językoznawczym. To również wiek intensywnego rozwoju technologii. Pojawiają się technologia EEG, tomografia komputerowa i rezonans magnetyczny.

Organizatorem pierwszego polskiego instytutu foniatry jako jednostki medycznej – Warszawskiego Instytutu Odchyleń Mowy – był Władysław Ołtuszewski, który zajmował się także popularyzacją wiedzy o zaburzeniach mowy. Głoszone przez niego tezy zawarte zostały w książce *O mowie i jej zaburzeniach* z 1906 roku. Podzielił zaburzenia na ośrodkowe: afazje (afazja percepcyjna, w głuchocie, w upośledzeniach umysłowych), dyzartrie (porażenia, jąkanie, rynolalia), oraz obwodowe – tylko dyzartrie.

Jednym z uznanych prekursorów problematyki zaburzeń mowy i języka w Polsce był psycholog i lekarz Mieczysław Choynowski. Jego prace z zakresu logopedii, w tym klasyfikacja zaburzeń mowy, miały istotny wpływ na rozwój tej dziedziny w Polsce. Napisał pracę pt. *Zarys logopedii czyli nauki o zaburzeniach mowy u dzieci* w roku 1913. Ta publikacja jest często uznawana za kamień milowy w rozwoju logopedii w Polsce. Choynowski wyróżnił kilka głównych rodzajów zaburzeń mowy, w tym: alalię rozumianą jako całkowity brak mowy u dzieci, dyslalię, którą określił jako trudności w prawidłowej artykulacji dźwięków

mowy, jąkanie – zaburzenie płynności mowy; załamki mowy; afazję, czyli utratę zdolności mówienia spowodowaną uszkodzeniem mózgu; bradyłalię – nadmiernie wolne tempo mowy; tachylalię – nadmiernie szybkie tempo mowy; rhinolalię – mowę nosową, spowodowaną nieprawidłowym przepływem powietrza przez jamę nosową i ustną. Choynowski zwracał uwagę na różnorodność zaburzeń mowy, jednak w klasyfikacji zaburzeń mowy kierował się głównie kryteriami objawowymi.

W powojennej Polsce problematyka zaburzeń mowy została podjęta w różnych ośrodkach w kraju – w Lublinie, Krakowie, Warszawie, Poznaniu. Za kolebkę polskiej nauki o zaburzeniach mowy niewątpliwie uznaje się ośrodek lubelski. Logopedia na Katolickim Uniwersytecie Lubelskim Jana Pawła II (KUL) została założona w 1962 roku. Stała się ona jedną z pierwszych akademickich jednostek tego typu w Polsce. Inicjatorką i założycielką logopedii na KUL była Halina Jastrzębowska, która przyczyniła się do rozwoju tej dyscypliny. W 1969 roku dzięki wiodącej roli uznanego za „ojca polskiej logopedii” – Leona Kaczmarka, na UMCS w Lublinie powstała Katedra Logopedii i Językoznawstwa Stosowanego. Również w Lublinie w 1963 roku z inicjatywy Leona Kaczmarka powstało Polskie Towarzystwo Logopedyczne. Do lat 80. XX w. istniały w Polsce cztery ośrodki kształcące logopedów. Oprócz wspomnianego ośrodka na UMCS należy wymienić Uniwersytet Warszawski – pierwotnie zakład prowadzony przez Halinę Mierzejewską. W 1974 powstało Podyplomowe Studium Logopedyczne prowadzone przez Irenę Styczek, a później między innymi przez Marię Przybysz-Piwkowską. Następnie założono ośrodki w Gdańsku (zakład utworzony przez Bronisława Rocławskiego) oraz na Uniwersytecie Gdańskim i w Opolu – Studium – firmowane przez Tadeusza Gałkowskiego i Grażynę Jastrzębowską. Obecnie kilka ośrodków naukowych zajmuje się badaniami języka w nurcie kognitywistycznym. Należą do nich między innymi Instytut Językoznawstwa na Uniwersytecie im. Adama Mickiewicza w Poznaniu. UAM ma silną tradycję badawczą w dziedzinie językoznawstwa kognitywnego, oferując różnorodne programy i projekty badawcze skoncentrowane na poznaniu języka i umysłu. Również Centrum Badań nad Kulturą, Językiem i Umysłem (Centrum Q) na Uniwersytecie Warszawskim. To interdyscyplinarne centrum koncentruje się na badaniach i edukacji w zakresie kognitywistyki, łącząc perspektywę z psychologii, informatyki, filozofii i językoznawstwa.

Ujęcie problemowe pojęcia

Problematyka zaburzeń mowy i języka traktowana była jako obszar zainteresowania w ramach różnych dyscyplin naukowych. W pierwszej kolejności trudności z komunikacją, używaniem języka podejmowane były w ramach nauk medycznych. W Międzynarodowej Statystycznej Klasyfikacji Chorób i Problemów Zdrowotnych ICD-10 zaburzenia mowy zostały wyodrębnione jako kategorie w ramach zaburzeń rozwoju psychologicznego i zostały z licznymi podkategoriami ujęte w kodzie F80 „Specyficzne zaburzenia rozwoju mowy i języka”.

Zaburzenia charakteryzujące się upośledzeniem normalnych wzorców nabywania umiejętności językowych od wczesnych stadiów rozwoju. Stanów tych nie można bezpośrednio wiązać z zaburzeniami neurologicznymi, nieprawidłowościami mechanizmów mowy, zaburzeń sfery czuciowej, upośledzeniem umysłowym lub czynnikami środowiskowymi. Specyficzne zaburzenia rozwoju mowy i języka często pociągają za sobą problemy, takie jak: trudności w czytaniu i ortografii, zaburzenie związków interpersonalnych oraz odchylenia w zakresie emocji i zachowania (ICD-10, 2008, s. 245).

Od 1 stycznia 2022 roku znana jest już klasyfikacja ICD-11, która nie obowiązuje jeszcze w Polsce (przewidziany jest pięcioletni tryb wdrażania nowej klasyfikacji).

W nurcie lingwistycznym, a dziś z całą pewnością kognitywistycznym, zaburzenia mowy traktujemy jako kompleksową kategorię problemów komunikacyjnych, które mogą wpływać na zdolność jednostki do wyrażania myśli oraz do skutecznego porozumiewania się z innymi. Wyodrębnione podstawowe funkcje języka, jakimi są reprezentatywna i komunikacyjna, wzajemnie się przenikają, co powoduje trudność z wyodrębnieniem ich pierwotnych znaczeń i wektorów wpływu. „Dzięki badaniom psychologicznym, neurologicznym, a nade wszystko lingwistycznym coraz jaśniejsze staje się, iż język jest nie tylko nośnikiem komunikacji interpersonalnej lecz także intrapersonalnej” (Jakobson, 1989, s. 60). W tym miejscu warto zwrócić uwagę, iż kognitywistyczne podejście pozwala na syntezę tych wyodrębnionych funkcji. Będąc użytkownikiem języka, stosujemy równocześnie i równoległe obie funkcje naraz jako narzędzia komunikowania się z innymi oraz jako narzędzia myślenia i samozwrotnego komunikowania się.

Rozpatrując w nurcie kognitywistycznym tak rozumianą funkcję komunikatywną, należy uwzględnić klasyfikację Romana Jakobsona. Zidentyfikował on sześć podstawowych funkcji języka i możemy równolegle dopasować sześć odpowiadających im zaburzeń mowy:

1. Funkcja referencyjna (denotatywna): Związana z kontekstem komunikatu. Zaburzenia tej funkcji manifestują się w trudnościach związanych z adekwatnym opisem rzeczywistości, co można obserwować np. w autyzmie lub afazji sensorycznej.
2. Funkcja poetycka: Skupia się na formie przekazu i jego aspekcie estetycznym. Zaburzenia tej funkcji objawiają się problemami z wykorzystaniem języka w sposób twórczy lub artystyczny, co jest typowe np. dla osób z zaburzeniami spektrum autyzmu.
3. Funkcja emotywna (ekspresywna): Wyraża postawę nadawcy do treści przekazu. Zaburzenia tej funkcji obejmują trudności w wyrażaniu emocji i stosunku do poruszanych tematów, widoczne np. w afazji motorycznej.
4. Funkcja fatyczna (kontaktowa): Dotyczy utrzymania komunikacji między nadawcą a odbiorcą. Zaburzenia tej funkcji obejmują trudności w nawiązywaniu lub utrzymywaniu rozmowy, co jest charakterystyczne dla osób z apatią lub z ciężkimi formami zaburzeń psychicznych.
5. Funkcja metajęzykowa: Umożliwia komunikację na temat samego języka. Zaburzenia tej funkcji manifestują się problemami z analizą i dyskusją na temat języka, np. w dysleksji lub pewnych typach afazji.
6. Funkcja konatywna (apelatywna): Związana z wpływem przekazu na odbiorcę, mająca na celu skłonienie go do określonego działania lub reakcji. Zaburzenia tej funkcji mogą objawiać się trudnościami w używaniu języka do przekonywania lub wydawania poleceń, co może być widoczne w różnych typach afazji (Jakobson, 1989).

Roman Jakobson odrębnie zajął się zakłóceniem, które nazywamy afazją, jako kwintesencją zaburzeń językowych. Będąc lingwistą, miał kompetencje do rozpatrywania języka we wszystkich jego stanach: języka w działaniu, języka w ruchu, języka w powstaniu i języka w rozkładzie (Jakobson, 1989, s. 150). Uznawał jego dwojaki charakter (w każdym

znaku językowym działają układy kombinacji i selekcji) i analizował językowe skutki schorzeń afatycznych na dwóch osiach: zakłócenie w dziedzinie podobieństwa oraz zakłócenie w dziedzinie przyległości. W dziedzinie podobieństwa, które jest jednym z dwóch głównych typów operacji językowych według Jakobsona (obok metafory i metonimii), zakłócenie manifestuje się przez użycie nieoczekiwanych porównań, metafor lub przez utrudnione przejście od wyrazu do jego synonimów i wyrażań opisowych, ale też heteronimów (co wiązałoby się z utratą zdolności poliglotycznych). Rzeczywistością językową chorego stałby się jego wewnętrzny „idiolekt”, język osobniczy wiązałoby się z utratą zdolności przechodzenia z jednego kodu (znaczeniowego) na drugi. Wówczas nadawca komunikatu nie byłby zdolny do dostrzeżenia związków między obiektami lub pojęciami, które są zwykle postrzegane jako podobne. W przypadku zakłóceń w dziedzinie przyległości następuje zatracenie zdolności tworzenia poprawnie skonstruowanych dłuższych wypowiedzi:

Ta afazja, polegająca na upośledzeniu zdolności do tworzenia kontekstu, a którą można by nazwać zakłóceniem w dziedzinie przyległości, zmniejsza zasięg i różnorodność zdań. Prawidła składniowe organizujące wyrazy w wyższe jednostki są utracone: ta utrata zwana agramatyzmem, zmienia zdanie w „kupę wyrazów” (Jakobson, 1989, s. 164).

Wyodrębnienie przez autora dwóch specyficznych rodzajów afatycznych zaburzeń, ale również innych, bezpośrednio wynikających z funkcji języka potwierdza pogląd Jakobsona na język:

Język jest systemem systemów, całościowym kodem, zawierającym różne poszczególne subkody. Różnorodne style mówienia nie tworzą przypadkowego, mechanicznego nagromadzenia, lecz prawidłową hierarchię poszczególnych kodów (Jakobson, 1989, s. 145).

Współcześnie koncepcja Jakobsona znalazłaby zastosowanie w programowaniu języka naturalnego przy pomocy sztucznej inteligencji (ang. *Natural Language Processing, NLP*). NLP to dziedzina sztucznej inteligencji, która zajmuje się interakcją między programami i algorytmami komputerowymi a ludzkim językiem. Zastosowanie koncepcji Jakobsona umożliwiłoby algorytmom rozumienie, interpretację i generowanie ludzkiego języka w sposób, który jest zarówno użyteczny, poprawny, jak i naturalny.

Interdyscyplinarne, kompleksowe i zintegrowane podejście do zrozumienia i leczenia zaburzeń mowy, uwzględniające zarówno neurobiologiczne, jak i poznawcze aspekty tych zaburzeń, oferowane jest na gruncie kognitywistyki. W kognitywistycznym ujęciu zaburzeń mowy kluczowe jest zrozumienie tego, jak różne procesy poznawcze i struktury mózgowe współpracują, aby umożliwić produkcję i zrozumienie języka. Przykładami kognitywnych modeli zaburzeń mowy są: model Wernicke-Lichtheim-Geschwind, model interakcyjny czy model sieci neuronowych.

Warto również zwrócić uwagę na behawioralną perspektywę w ujmowaniu zaburzeń mowy, która zaprezentowana została już w latach 60. ubiegłego wieku przez polskiego badacza – Leona Kaczmarka. W duchu strukturalistycznym traktował mowę jako składową poszczególnych segmentów występujących niejako autonomicznie. Wobec tego realizacja każdego z wyodrębnionych przez niego komponentów mogła generować odrębne trudności. I tak spreparowane trzy części składowe: treść, forma językowa i substancja foniczna, miały charakteryzować się specyficznymi dla siebie przejawami zaburzeń. Jako kryterium uwzględnił objawowy ich charakter. W zakresie zaburzenia treści wyodrębnił: zaburzenia procesu uogólniania i abstrakcji, brak logiki w tekstach, nieprawidłowości w ukierunkowaniu myślenia; w zakresie zaburzenia języka: niemotę, afazję, alalię, agramatyzm, przejęzyczenie, a zaburzenia substancji w płaszczyźnie suprasegmentalnej: jąkanie, rynolalię, wymowę bezkrtaniowców, afonię, gielkot, bradylalię; w płaszczyźnie segmentalnej: dyslalię i w płaszczyźnie segmentalnej i suprasegmentalnej: palatolalię i mutyzm (Kaczmarek, 1975). Krytyczną analizę propozycji Leona Kaczmarka przedstawił Stanisław Grabias (2020). Za niemożliwe do obronienia uznał postrzeganie komponentów języka jako niezależnych od siebie, następujących sekwencyjnie po sobie aktów mowy (najpierw myślimy, potem myśl ubieramy w formę językową, a dopiero później nadajemy jej substancję). Jak przekonywał, mózg przecież pracuje symultanicznie. Za nieprawdziwe uznał również niektóre szczegółowe twierdzenia autora klasyfikacji dotyczące zaburzeń mowy w chorobach psychicznych czy jąkania. Grabias docenił natomiast udaną próbę uporządkowania terminologicznego i przełamanie biologizmu obowiązującego w objawowych klasyfikacjach.

Powołując się na rekomendację Międzynarodowego Towarzystwa Logopedów i Foniatorów oraz Europejskiej Unii Foniatorów, Tomasz

Zalewski wyodrębnił zaburzenia mowy w kontekście jej rozwoju, czyli: specyficzne opóźnienia rozwoju mowy, które nie nosiłoby znamion patologii oraz zaburzenia rozwoju mowy, traktowane jako odstępstwo od normy (w wyniku zaburzeń słuchu, centralnych ośrodków mowy, upośledzenia umysłowego, mózgowego porażenia dziecięcego oraz na tle emocjonalnym) (Zalewski, 1993, s. 48). Można uznać, że autor dostrzega kognitywny charakter zaburzeń mowy. Po pierwsze, dlatego, że według autora droga do umysłu wiedzie przez słuch i procesy słuchania (w tym przetwarzania słuchowego na poziomie neuronalnym), po drugie, jego zdaniem mowa ma podłoże neuronalne (zwraca uwagę na towarzyszące odstępstwa od normy) i w końcu, po trzecie, autor odwołuje się do mentalnych aspektów języka – zaburzeń zachowania, emocjonalnych i upośledzenia umysłowego.

Dla autorki pierwszej w Polsce monografii poświęconej logopedii Ireny Styczek to etiologia stała się najważniejszym kryterium klasyfikowania trudności z realizacją mowy i języka. I tak Styczek wyodrębniła dwie główne kategorie: zaburzenia zewnątrzpochodne (egzogenne, środowiskowe, które autorka uznaje za zaburzenia, a nie wadę) i zaburzenia wewnątrzpochodne (endogenne, właściwe, tożsame z wadą mowy). Wszystkie wewnątrzpochodne byłyby zaburzeniami mowy, które określa jako wady mowy. Owe *stricto* wady mowy w zależności od przyczyny według badaczki obejmują: dysglosję – zniekształcenie dźwięków mowy lub niemożność ich wytwarzania na skutek nieprawidłowej budowy narządów mowy lub obniżenia słyszalności; dysartrię (anartria) – zniekształcenie dźwięków mowy lub niemożność ich wytwarzania na skutek uszkodzenia ośrodków i dróg unerwiających narządy mowne; dyslalię (alalia) – opóźnienie w przyswajaniu sobie języka na skutek opóźnionego wykształcenia się funkcji pewnych struktur mózgowych; afazję – utratę częściową lub całkowitą znajomości języka na skutek uszkodzenia pewnych struktur mózgowych; jąkanie – zaburzenie płynności mowy (rytmu i tempa), którego przyczyna nie została ustalona; nerwice mowy (logoneurozy) – mutyzm, afonia, jąkanie, zaburzenia tempa mowy, modulacji siły i wysokości głosu u osób cierpiących na nerwice; oligofazję – niedokształcenie mowy spowodowane upośledzeniem umysłowym; schizofazję – mowę u osób chorych psychicznie i z zaburzeniami myślenia (Styczek, 1983, s. 250–251). Dziś trudno uznać tę klasyfikację za wystarczającą. Podstawowe kategorie

zaburzeń mowy zaproponowane przez Styczek (np. dyslalia, afazja, dysfazja) nadal są używane, ale często rozszerza się je i modyfikuje dzięki nowym technologiom diagnostycznym i terapeutycznym (np. neuroobrazowaniu, sztucznej inteligencji). Wartość prac Styczek polega na tym, że stanowią one solidny fundament, na którym można budować i rozwijać nowoczesne podejścia terapeutyczne.

Perspektywę lingwistyczną prezentował Stanisław Kania. Zawęził on rozumienie zaburzenia mowy głównie do aspektu formalnego. Według autora forma i struktura wypowiedzi (znaków językowych) stanowiłaby meritum zaburzenia. Jeśli tak, to poszczególne elementy wypowiedzi, w różnych ich konfiguracjach począwszy od płaszczyzny izolowanych głosek, poprzez sylaby, wyrazy aż do całej wypowiedzi stanowiłyby segmentalny aspekt zaburzeń. Z kolei zaburzenie segmentów wypowiedzi mogłoby przejawiać się na poziomie inwentarza głosek – paradygmatycznym, oraz wewnątrz struktury fonetycznej głosek – syntagmatycznym. Kania uważa, że zaburzenia syntagmatyczne i paradygmatyczne rezonują ze sobą wzajemnie. Mogą więc występować w postaci wtórnej i pierwotnej. Tak więc, jak pisze, „u tego samego pacjenta mogą istnieć: 1) pierwotne zaburzenia paradygmatyczne, 2) pierwotne zaburzenia syntagmatyczne, 3) wtórne zaburzenia paradygmatyczne, 4) wtórne zaburzenia syntagmatyczne” (Kania, 1982, s. 10). Przy czym wszystkie z kategorii wtórnych uznaje za zdeterminowane przez strukturę danego języka lub indywidualną realizację wymawianiową, co byłoby według autora równoznaczne z przyczynami wewnętrznymi. Proponując własną, autorską klasyfikację pierwotnych zaburzeń paradygmatycznych wyprowadza wywód z trójczłonowego podziału dokonanego przez foniatrów (mogilalię, paralalię i deformację). W jej miejsce wprowadza swój podział na elizję (mogilalia), substytucję (paralalia) i – finalnie – deformację. Strukturalny punkt widzenia autor konsekwentnie uwiadcza w proponowanym podziale zaburzeń syntagmatycznych. Mogą mieć charakter ilościowy, jakościowy, sekwencyjny, kombinowany i niesystemowy. Dla każdego z przejawów zaburzeń precyzyjnie dookreślił podkategorie.

Proponowana przez Kanię klasyfikacja zawęża możliwość eksploatacji zaburzeń do czysto formalnego aspektu. Jednak dzięki prezentowanym przez niego modelom z zakresu analizy języka naturalnego do badania zaburzeń mowy możemy tworzyć modele kognitywne, które

lepiej odzwierciedlają rzeczywiste mechanizmy przetwarzania języka, a dzięki technologiom (za pomocą narzędzi AI) mogą automatyzować procesy diagnostyczne.

Rozbudowaną próbę ujęcia różnorodności form patologii mowy podjęły autorki graficznej prezentacji przyczyn i objawów Halina Mierzejewska i Danuta Emiluta-Rozya. Sama klasyfikacja ma swoją historię i ewoluuje już od wielu lat (pierwsza propozycja pojawiła się w 1997 roku). Mimo że heterogeniczność konwergencji i dywergencji zaburzeń mowy nigdy nie została wyrażona przez autorki *explicite*, to niewątpliwie w tworzeniu kategorii form zaburzeń miało to zdecydowane znaczenie. Heterogeniczność dywergencji zaburzeń odnosi się do sytuacji, w której różnorodne elementy (przyczyny i objawy patologii) rozbiegają się w różnych kierunkach lub wykazują różnorodność w swoich trajektoriach zmian. Heterogeniczność konwergencji to sytuacja lub proces, w którym różnorodne elementy (przyczyny zaburzeń lub objawy) zbiegają się lub łączą się mimo swojej pierwotnej różnorodności. Jak piszą autorki,

analiza graficznej formy naszego zestawienia pokazuje (...), że ten sam zakres objawów może być wywołany przez różne przyczyny endogenne, egzogenne, lub endo-egzogenne (Emiluta-Rozya, 2012, s. 76).

Pokazując najszersze możliwe pole zakresów objawów i przyczyn, jednocześnie przynajmą:

nie znalazłyśmy sposobu na obrazowe zróżnicowanie zaburzeń czynności rozumienia i tworzenia wypowiedzi słownych w poszczególnych formach patologii mowy (Emiluta-Rozya, 2012, s. 75).

Wobec tego, ze względów funkcjonalnych, autorki zaznaczają, że proponowana typologia może być użyteczna przede wszystkim dla praktyków – logopedów. Proponowana przez nie rozbudowana grafika niczym suwak logarytmiczny pozwalałaby na poruszanie się w obrębie możliwych przyczyn, form i objawów. Mogłaby też stanowić podłoże dla programowania logarytmów ułatwiających budowanie „sita” informacyjnego dla różnych indywidualnych wypowiedzi.

Przyjęta przez Stanisława Grabiasa (2020) perspektywa, która obejmuje aspekty poznawcze oraz neurobiologiczne w ich ścisłych

wzajemnych relacjach, doprowadziła go do własnej klasyfikacji zaburzeń mowy. Według autora, po pierwsze, język przynależy tylko do człowieka dzięki złożonej strukturze, jaką jest mózg ludzki, po drugie, do pewnego wieku dzieci nabywają język samorzutnie i dzięki ścieżkom neuronalnym interioryzują go do szóstego roku życia, potem następuje nabywanie języka na drodze programowania i ostatnie przyjęte przez autora założenie – nabywamy język poprzez słuch i tylko on jest gwarantem właściwej drogi języka do umysłu. Z przesłanek tych wyprowadza trzy istotne założenia badawcze będące podstawą szczegółowo prezentowanej klasyfikacji: 1) wypowiedzi – ich znaczenie i struktura – stanowią odzwierciedlenie sprawności interakcyjnych: systemowej, społecznej, sytuacyjnej i pragmatycznej; 2) poza tym istnieją odmienne kompetencje językowe i komunikacyjne (samorzutne byty w umysłach organizujące zdania oraz wypowiedzi dialogowe i narracyjne); 3) sprawności formacyjne wyrażające się w biologicznych procesach warunkujących język obejmują takie jego aspekty jak: słuch fizyczny, przetwarzanie słuchowe na poziomie neuronalnym ze słuchem fonemowym, plastyczność mózgu i układy kostne z mięśniowymi. Zaburzenia mowy byłyby więc konceptualnym konglomeratem przyczyn dysfunkcji, ich skutków, a różnorodność zaburzeń warunkowałyby również antycypacyjny charakter działań naprawczych. Wobec tak prezentowanego konceptu języka i jego udziału w formowaniu problematyki zaburzeń mowy Stanisław Grabias zaproponował następującą klasyfikację:

1. Zaburzenia mowy ujawniające się na tle niewykształconych sprawności biologicznych warunkujących język: głuchota i niedosłuch, alalia i dyslalia (kompetencje nie wykształcają się w związku z centralnymi zaburzeniami przetwarzania słuchowego), oligofazja, autyzm, padaczki dziecięce, płodowy zespół alkoholowy.
2. Zaburzenia mowy ujawniające się na tle niedostatecznego wykształcenia sprawności realizacyjnych: dysglosja, gielkot, jąkanie, dyzartria (związana jest z uszkodzeniami centralnego lub obwodowego układu nerwowego).
3. Zaburzenia mowy ujawniające się na tle rozpadu możliwości językowych i komunikacyjnych: afazja (związana z utratą zdolności rozumienia i tworzenia wypowiedzi na skutek uszkodzenia struktur mózgowych, a trudności wiążą się z zaburzeniami

mechanizmów programujących czynności mowy), pragnozja (zaburzenia mowy powstałe na skutek uszkodzenia „niejęzykowej” pólki mózgu i ujawniające się w postaci trudności w rozumieniu wyrażeń metaforycznych, dłuższych wypowiedzi oraz w rozpoznawaniu emocjonalnych stanów mówcy), schizofazja (jest skutkiem schizofrenii), demencja (jako efekt zamierania neuronów i układów synaptycznych, ujawnia się np. w chorobie Alzheimera) (Grabias, 2020).

Koncepcja zaburzeń mowy prezentowana przez Stanisława Grabiasa z całą pewnością lokuje się w interdyscyplinarnym nurcie badawczym, jakim jest kognitywistyka. Jego podejście zakłada, że zaburzenia mowy nie są jedynie wynikiem problemów na poziomie wykonawczym czy anatomicznym, ale często wynikają z zakłóceń w procesach programujących mowę, takich jak pamięć, uwaga, percepcja czy myślenie. W swoich pracach Grabias podkreślał znaczenie procesów poznawczych w komunikacji językowej i zaburzeniach mowy. Uwzględniał szeroki kontekst funkcjonowania człowieka, w tym aspekty poznawcze, emocjonalne i społeczne, które wpływają na zdolność komunikacyjną.

Refleksja systematyczna z wnioskami i rekomendacjami

Język jest postrzegany jako integralna część ludzkiego procesu poznawczego, a więc jest rozumiany nie tylko jako system komunikacji, ale przede wszystkim jako narzędzie umożliwiające myślenie i organizację wiedzy. Struktury językowe są reprezentowane w umyśle i są używane w procesach myślenia i rozumienia świata. Wobec tego zaburzenia mowy i języka stanowią złożoną kategorię problemów komunikacyjnych, które wymagają interdyscyplinarnego podejścia zarówno w diagnozie, jak i terapii. Stanisław Grabias podkreśla, że tezy wynikające z funkcji, jakie pełni język mają decydujące znaczenie w wyznaczaniu procesu naprawczego, jakim jest oddziaływanie terapeutyczne. Szczegółowa wiwisekcja z gruntownym, szczegółowym analizowaniem zachowań językowych ujawnia obszary oddziaływania i ukierunkowuje to oddziaływanie. To, że język jest drogą do poznawczych aspektów człowieka (wiedzy o sobie, świecie, umiejętności korzystania z niej),

jest narzędziem wyrażania ludzkich emocji i pragnień oraz osadza jednostkę w życiu społecznym, wyznacza azymut w ramach modelu standardowych procedur (Grabias, 2020). Tak rozumiany kognitywny aspekt języka znajduje swoje pragmatyczne zastosowanie w diagnozie i terapii.

Rozwój metod diagnostycznych i terapeutycznych jest kluczowy dla poprawy jakości życia osób z zaburzeniami komunikacyjnymi. Diagnoza opierać się będzie na kompleksowej ocenie, która może obejmować badania medyczne, ocenę mowy i języka, a także testy psychologiczne. Natomiast terapia zaburzeń powinna uwzględniać różnorodne podejścia w zależności od rodzaju i ciężkości zaburzenia. Metody te mogą obejmować terapię mowy i języka, terapię behawioralną, a także wykorzystanie technik wspomagających komunikację (tzw. AAC – *Augmentative and Alternative Communication*).

Ponieważ kognitywistyka jako nauka interdyscyplinarna interesuje się między innymi mechanizmami przetwarzania języka – zarówno w zakresie produkcji (mówienie i związane z tym zaburzenia), jak i rozumienia (słuchanie) – wiedzę o zaburzeniach mowy i języka można wykorzystać w badaniach nad komputerowymi modelami przetwarzania języka naturalnego oraz w eksperymentach psycholingwistycznych.

BIBLIOGRAFIA

- Emiluta-Rozya, D. (2012). Formy zaburzeń mowy. W: S. Grabias, & M. Kurkowski (Red.), *Logopedia. Teoria zaburzeń mowy. Podręcznik akademicki* (s. 73–87). Lublin: Wydawnictwo UMCS.
- Grabias, R. (2012). Teoria zaburzeń mowy. Perspektywy badań, typologie zaburzeń, procedury postępowania logopedycznego. W: S. Grabias, & M. Kurkowski (Red.), *Logopedia. Teoria zaburzeń mowy. Podręcznik akademicki* (s. 15–71). Lublin: Wydawnictwo UMCS.
- Grabias, S. (2020). Logopedia polska w stuleciu Polski Niepodległej. Trzy perspektywy rozwoju dyscypliny. *Logopedia*, 47(2), 517–537.
- Greene, J. (1977). *Psycholingwistyka. Chomsky a psychologia*. Warszawa: PWN.
- ICD-10 (2008). Międzynarodowa Statystyczna Klasyfikacja Chorób i Problemów Zdrowotnych, rewizja dziesiąta. T. I. World Health Organization.
- Jakobson, R. (1989). *W poszukiwaniu istoty języka 1, 2. Wybór pism*, M.R. Mayenowa (Red.). Warszawa: Państwowy Instytut Wydawniczy.

- Kaczmarek, L. (1975). Korelacyjna klasyfikacja zaburzeń słownego i pisemnego porozumiewania się. *Logopedia*, 12, 5–13.
- Kania, J.T. (1982). *Szkice logopedyczne*. Warszawa: Wydawnictwo Szkolne i Pedagogiczne.
- Kurcz, I. (1987). *Język a reprezentacja świata w umyśle*. Warszawa: Wydawnictwo PWN.
- Mayenowa, M.R. (1989). Roman Jakobson – uczoney i człowiek. W: R. Jakobson, *W poszukiwaniu istoty języka*. Warszawa: Państwowy Instytut Wydawniczy.
- Styczek, I. (1983). *Logopedia*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Zalewski, T. (1993). Klasyfikacja zaburzeń mowy. W: T. Gałkowski, Z. Tarkowski, & T. Zalewski (Red.) *Diagnoza i terapia zaburzeń mowy* (s. 7–39). Lublin: Wydawnictwo UMCS.

Sławomir Sztajer

Uniwersytet im. A. Mickiewicza w Poznaniu

<https://orcid.org/0000-0001-8975-2994>

<https://doi.org/10.35765/slowniki.423>

Religioznawstwo kognitywne

Streszczenie

DEFINICJA POJĘCIA: Religioznawstwo kognitywne jest interdyscyplinarnym podejściem w badaniach nad religiami zajmującym się poznawczymi uwarunkowaniami myślenia i zachowania. Bada procesy i mechanizmy poznawcze, zaangażowane w tworzenie, przyswajanie i przekazywanie idei religijnych, a także procesy i mechanizmy leżące u podłoża zachowania rytualnego. Podejście kognitywne sięga także do wiedzy o ewolucji ludzkiego umysłu, wiedzy o neuronalnych podstawach doświadczenia religijnego, a także badań nad interakcją umysłu z otaczającym światem.

ANALIZA HISTORYCZNA POJĘCIA: Systematyczne kognitywne badania nad religią zostały zapoczątkowane w latach 90. ubiegłego stulecia. Sformułowano wówczas szereg teorii, które wyznaczyły kierunki badań w kolejnych dziesięcioleciach. Rozwojowi religioznawstwa kognitywnego towarzyszyło niezadowolenie z osiągnięć dotychczasowego religioznawstwa.

UJĘCIE PROBLEMOWE POJĘCIA: Religioznawstwo kognitywne przyjmuje, że religia jest rezultatem działania naturalnych, zwyczajnych mechanizmów poznawczych. Religie raczej opierają się na mechanizmach umysłowych, które wyewoluowały w odpowiedzi na problemy, jakie przez miliony lat napotykali w środowisku nasi przodkowie. Teza o naturalności religii ma ważne konsekwencje metodologiczne: zjawiska religijne należy badać tak, jak bada się wszystkie inne zjawiska naturalne.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Podstawową zaletą podejścia kognitywnego w badaniach nad religiami jest

jego interdyscyplinarność oraz dążenie do uwzględnienia różnych warunkowań, jakim podlegają przekonania i zachowania religijne, a także integracji wiedzy z zakresu kognitywistyki, neuronauki, teorii ewolucji i nauk o kulturze.

Słowa kluczowe: religia, rytuał, procesy poznawcze, ewolucja

Definicja pojęcia

Kognitywne nauki o religii stanowią stosunkowo nowy obszar badawczy, który powstał w ramach szeroko pojętej kognitywistyki w latach 90. XX wieku. Współczesna kognitywistyka religii stanowi wielodyscyplinarne, a niekiedy również interdyscyplinarne podejście w badaniach nad religiami. Rozwijana jest nie tylko przez kognitywistów, ale również przez religioznawców, antropologów, psychologów kognitywnych, ewolucyjnych, społecznych i rozwojowych, historyków, archeologów, kulturoznawców, a także biologów (White, 2021, s. 1). Termin *cognitive science of religion* można przełożyć jako „kognitywistyka religii” i takie tłumaczenie tego terminu byłoby uzasadnione, biorąc pod uwagę genezę tego podejścia. Kognitywistyka religii stanowiłaby wówczas subdyscyplinę nauk kognitywnych. Biorąc jednak pod uwagę coraz większą rolę podejścia kognitywnego w religioznawstwie, można również mówić o religioznawstwie kognitywnym, podobnie jak mówi się o antropologii czy socjologii kognitywnej. Religioznawstwo kognitywne może być rozumiane bądź jako subdyscyplina religioznawstwa, bądź orientacja teoretyczno-metodologiczna stosowana w naukach o religii. Występowanie dwóch różnych sposobów rozumienia kognitywnych badań nad religią nie jest niczym niezwykłym, ponieważ także inne subdyscypliny współczesnego religioznawstwa, takie jak psychologia, socjologia czy antropologia religii, funkcjonują równocześnie jako subdyscypliny odpowiednich nauk – psychologii, socjologii i antropologii.

Kognitywne podejście w badaniach nad religiami opiera się na założeniu, że myślenie i zachowanie religijne mogą być wyjaśniane na podstawie naturalnych procesów i mechanizmów, które składają się na ludzki system poznawczy. Leżące u podłoża religii procesy i mechanizmy poznawcze są wyjaśniane w kategoriach ewolucyjnych. Kognitywni badacze religii interesują się zatem nie tylko tym, jak działanie ludzkiego umysłu wpływa na religię, ale także w jaki sposób umysł został ukształtowany na drodze selekcji naturalnej w środowisku zamieszkiwanym przez naszych przodków. Argumentuje się, że przynajmniej niektóre mechanizmy umysłowe zostały ukształtowane w odpowiedzi na powtarzalne problemy, z jakimi człowiek mierzył się w ewolucyjnej przeszłości. Perspektywa ewolucyjna, zwłaszcza zaś psychologia ewolucyjna, od początku była obecna w kognitywnych badaniach nad religią (Liddle

& Shackelford, 2021). Badania nad umysłem, do których odwołują się kognitywni badacze religii, obejmują także poszukiwanie neuronalnych korelatów myślenia i zachowania religijnego oraz złożonych interakcji umysłu ze środowiskiem przyrodniczym i społecznym (Schjoedt, 2009).

Punktem wyjścia kognitywnych badań nad religią były pytania dotyczące przyswajania, reprezentowania oraz przetwarzania pojęć i przekonań religijnych, a także pytania dotyczące wpływu reprezentacji religijnych na zachowanie. Kognitywiści poszukują ponadto odpowiedzi na fundamentalne kwestie, takie jak geneza religii, powszechne występowanie idei i zachowań religijnych w kulturach świata, jak również łatwość, z jaką idee religijne są przyswajane przez jednostki ludzkie i ulegają rozpowszechnieniu w populacjach ludzkich. Kognitywne badania nad religiami obejmują kluczowe wymiary religii: pojęcia i przekonania religijne, działania rytualne, doświadczenie religijne, formy społecznej organizacji życia religijnego, etyki religijne. Przedmiotem zainteresowania są ponadto zjawiska tradycyjnie badane przez religioznawców, choć niebędące konstytutywnymi elementami tradycji religijnych, takie jak magia czy mit.

Kognitywne badania nad religią mają charakter wielodyscyplinarny. Uczestniczą w nich nie tylko religioznawcy, ale także specjaliści z zakresu psychologii, antropologii, socjologii, neuronauki, biologii, historii czy archeologii. Nie chodzi przy tym jedynie o zestawienie osiągnięć wymienionych nauk, lecz o tworzenie projektów interdyscyplinarnych, zmierzających ku integracji wiedzy wytworzonej w różnych dyscyplinach i uzyskaniu kompleksowego rozumienia problemów badawczych. Szeroka interdyscyplinarność zapobiega redukowaniu zjawisk religijnych do zjawisk jednego typu – zabiegu, który w dziejach religioznawstwa prowadził do konstruowania uproszczonych, jednowymiarowych teorii religii. Zamiast tego kognitywiści konstruują wieloprzyczynowe modele i uwzględniają różne aspekty zjawisk.

Analiza historyczna pojęcia

Systematyczne kognitywne badania nad religią zostały zapoczątkowane w latach 90. ubiegłego stulecia, jednak pierwsza praca zwiastująca rozwój tego nurtu badań została opublikowana w 1980 roku. Był

to artykuł Stewarta Guthrie'go zatytułowany *Kognitywna teoria religii*, w którym autor zaproponował kognitywne podejście do wierzeń i praktyk religijnych, wskazując na antropomorfizację jako proces odgrywający istotną rolę w myśleniu religijnym. Kilkanaście lat później teoria Guthrie'go została rozbudowana i zaprezentowana w formie monograficznej (Guthrie, 1993). W latach 90. zaproponowano kilka innych kognitywnych teorii religii, które wyznaczyły główne kierunki badań. Do najważniejszych należały: kognitywna teoria religijnych systemów rytualnych (Lawson & McCauley, 1990); teoria kulturowej transmisji idei religijnych Pascala Boyera, podkreślająca rolę ograniczeń poznawczych w powstawaniu i rozpowszechnianiu idei religijnych (Boyer, 1994); prace Harveya Whitehouse'a, w których zaprezentowana została wczesna wersja teorii dwóch stylów religijności, obrazowego i doktrynalnego, rozwinięta w późniejszych opracowaniach monograficznych (Whitehouse, 2004). W kolejnych dekadach sformułowano szereg nowych hipotez, z których wiele poddano testom empirycznym, w tym badaniom eksperymentalnym.

Rozwojowi kognitywistyki religii towarzyszyły niezadowolenie z osiągnięć dotychczasowego religioznawstwa, w którym istotną rolę odgrywały podejścia fenomenologiczne, historyczno-porównawcze, społeczno-kulturowe, a także wątpliwości dotyczące metod stosowanych w badaniach religioznawczych. Dostrzeżono na przykład, że w badaniach tych dominuje podejście interpretacyjne, natomiast niewiele mówi się o wyjaśnianiu zjawisk religijnych. Stwierdzono, że sposobem na przewyżczenie tego ograniczenia dotychczasowego religioznawstwa nie jest zastąpienie interpretacji wyjaśnieniem, ale przywrócenie między nimi równowagi poprzez konstruowanie teorii wyjaśniających (Lawson & McCauley, 1990, s. 14). Niektóre podejścia interpretacyjne traktowały wyjaśnianie zjawisk religijnych jako próbę ich redukcji do zjawisk innego typu. Zwłaszcza badacze należący do nurtu fenomenologicznego postulowali ujmowanie religii jako rzeczywistości *sui generis*, argumentując, że zjawiska religijne są nieredukowalne do zjawisk niereligijnych badanych przez inne nauki. Podejście kognitywne kwestionuje przekonanie o autonomii i nieredukowalności religii. Zjawiska religijne są częścią ludzkiego świata i podlegają także naturalnym uwarunkowaniom. Jako takie mogą być wyjaśniane na bazie zjawisk niereligijnych.

Przedmiotem krytyki kognitywistów były także różne podejścia społeczno-kulturowe, zwłaszcza te, które akcentowały zróżnicowanie

i nieporównywalność zjawisk religijnych, jak również eksponowały kulturowe determinanty ludzkiego zachowania. Zdaniem kognitywistów w charakteryzujących się olbrzymim zróżnicowaniem zjawiskach religijnych dostrzec można powtarzalne idee i zachowania, które można wyjaśnić w kategoriach uniwersalnej struktury ludzkiego umysłu. Czynniki kulturowe mają istotne znaczenie, ale nie sposób różnych form życia religijnego wyjaśnić jedynie wpływem kultury. Co więcej, dynamika systemów kulturowych w znacznym stopniu kształtowana jest przez właściwe ludzkim umysłom ograniczenia poznawcze związane z funkcjonowaniem pamięci, uwagi i przetwarzaniem informacji. Ograniczenia te z kolei można wyjaśniać na podstawie ugruntowanych teorii nauk biologicznych, czego przykładem jest zastosowanie teorii ewolucji oraz wiedzy o mózgu.

W podejściu kognitywnym zwrócono także uwagę na te aspekty funkcjonowania religii, które wcześniej rzadko były przedmiotem badań. Dostrzeżono, że zarówno idee religijne, jak i przekonania ludzi wierzących nie ograniczają się do idei i przekonań składających się na doktryny poszczególnych religii. Oficjalna doktryna może mieć charakter normatywny, określać, w co wierzyć należy, ale z pewnością nie wyczerpuje zasobu idei, którymi posługują się ludzie religijni. Religioznawstwo, zdaniem kognitywistów, przywiązywało zbyt dużą wagę do oficjalnych doktryn religijnych i deklarowanych przekonań, mniej uwagi natomiast poświęcało temu, jak ludzie religijni rzeczywiście myślą. Metody eksperymentalne umożliwiające dotarcie do tej „nieoficjalnej” wersji myślenia religijnego są często stosowane w kognitywnych badaniach religii.

Ujęcie problemowe pojęcia

Fundamentalną tezę stawianą w ramach podejścia kognitywnego jest przekonanie o naturalności religii. Religia jest zjawiskiem naturalnym nie tylko w tym sensie, że zależy od uniwersalnych, pozakulturowych uwarunkowań poznawczych, ale także dlatego, że jest rezultatem działania naturalnych, zwyczajnych mechanizmów poznawczych (Boyer, 1994). U podłoża myślenia religijnego nie leżą żadne specyficzne dla religii zdolności poznawcze; religie raczej opierają się na mechanizmach umysłowych, które wyewoluowały w odpowiedzi na problemy,

jakie przez miliony lat napotykali w środowisku nasi przodkowie. Teza o naturalności religii ma ważne konsekwencje metodologiczne: zjawiska religijne należy badać tak, jak bada się wszystkie inne zjawiska naturalne. Wyjaśnienie religii w szczególności nie wymaga postulowania istnienia jakiegokolwiek nadprzyrodzonej rzeczywistości.

Znaczna część kognitywnych badaczy religii stosuje podejście ewolucyjne w wyjaśnianiu zjawisk religijnych. Przyjmuje się, że umysł stanowi wytwór ewolucji i jest złożony z wyspecjalizowanych mechanizmów poznawczych, z których każdy stanowi odpowiedź na określony problem adaptacyjny, z którym mierzyli się nasi przodkowie. Mechanizmy poznawcze leżące u podłoża wierzeń i przekonań religijnych wyjaśniane są na podstawie teorii ewolucji. Wśród badaczy zajmujących się ewolucyjnymi podstawami religii nie ma zgody co do tego, czym są idee i zachowania religijne w świetle teorii ewolucji. Zwolennicy adaptacjonizmu utrzymują, że przynajmniej niektóre wierzenia i zachowania religijne mają charakter adaptacyjny w tym znaczeniu, że zwiększają szansę na przeżycie i reprodukcję jednostek lub grup ludzkich, a także zwiększają dostosowanie jednostek lub grup w stosunku do innych jednostek lub grup ludzkich. Za adaptacyjnym charakterem przynajmniej niektórych religii przemawia fakt, że wspierają one radzenie sobie z różnymi problemami adaptacyjnymi, m.in. mogą mieć pozytywny wpływ na zdrowie fizyczne i psychiczne, dostarczać jednostkom sensownych, całościowych i optymistycznych wizji świata, ograniczać racjonalne motywy oszukiwania innych i tym samym stabilizować wymianę społeczną, podtrzymywać kooperację między jednostkami należącymi do tej samej grupy społecznej. Z drugiej strony – i jest to argument przeciw adaptacyjnemu charakterowi religii – zachowania religijne wiążą się z istotnymi kosztami ponoszonymi przez wyznawców (czas, energia, zasoby materialne), a w niektórych przypadkach stanowią zagrożenie dla życia i dobrostanu jednostek i grup społecznych (skrajna asceza, samobójstwa rytualne, konflikty na tle religijnym).

Większość badaczy kognitywnych studiujących religie kwestionuje adaptacyjny charakter zachowań religijnych. Argumentują oni, że wierzenia i zachowania religijne są raczej produktem ubocznym adaptacyjnych mechanizmów poznawczych, ale same nie posiadają wartości adaptacyjnej. Przykładowo pojęcie Boga jako istoty posiadającej umysł, myślącej, czującej i działającej w sposób intencjonalny nie jest

wprawdzie adaptacją, ale wykorzystuje będący adaptacją mechanizm poznawczy leżący u podłoża ludzkiej zdolności do reprezentowania stanów umysłowych innych osób i przewidywania ich zachowania. Adaptacja ta wyewoluowała w odpowiedzi na złożoność życia zbiorowego, niezależnie od religii. Nieobce podejściu kognitywnemu są ponadto próby wyjaśniania religii w kategoriach ewolucji kulturowej. W tym przypadku jednostką ewolucji jest pewna cecha kulturowa lub mem.

W badaniach nad religiami znaczną wagę przywiązywano do przekonań, które składają się na doktrynę danej religii, są werbalizowane, systematyzowane i racjonalizowane przez ekspertów religijnych. Przekonania takie składają się na oficjalną doktrynę i z pewnością mają istotny wpływ na myślenie wyznawców. Jednak – jak zauważają badacze kognitywni – myślenie religijne nie ogranicza się do pojęć i przekonań doktrynalnych, ponieważ w życiu codziennym ludzie religijni mają skłonność do posługiwania się ideami religijnymi znacznie odbiegającymi od oficjalnej doktryny. Zdaniem Justina L. Barretta (2004) zachowanie ludzi religijnych wykazuje wyraźną skłonność do swoistej poprawności teologicznej – deklarowania przekonań zgodnych z oficjalną doktryną, a jednocześnie posługiwania się pojęciami i przekonaniem znacznie odbiegającymi od doktrynalnej normy w sytuacjach, które wymagają szybkiego przetwarzania informacji. Przykładowo w trakcie modlitwy osoba religijna może przedstawiać sobie Boga nie jako istotę wszechwiedzącą, wszechobecną, niezmienną oraz transcendentną, ale jako bardzo podobną do człowieka osobę rozumiejącą język etniczny, w jakim wierzący się modli, a także poddaną różnym ograniczeniom: czasowym, przestrzennym, poznawczym, emocjonalnym. Pomimo że w niektórych religiach doktryna nakazuje myśleć o Bogu jako istocie fundamentalnie różnej od człowieka, w rozmaitych sytuacjach życia codziennego wierzący posługują się antropomorficznym wyobrażeniem Boga.

Stosowanie różnych pojęć istot nadprzyrodzonych jest konsekwencją istnienia rozmaitych systemów przetwarzania informacji przez ludzki umysł. Wśród badaczy nie ma pełnej zgody, jak należy te systemy różnicować. Niektórzy badacze postulują istnienie dwóch systemów przetwarzania. Dla uproszczenia przyjąć można, że umysł funkcjonuje, opierając się na dwóch systemach przetwarzania: intuicyjnym i refleksyjnym. Przetwarzanie intuicyjne jest szybkie, automatyczne i nieświadome, natomiast przetwarzanie refleksyjne jest wolne, kontrolowane

i świadome. Przetwarzanie informacji bazujące na tych dwóch różnych systemach nie jest specyficzne dla myślenia religijnego, ale w religiach odgrywa szczególnie istotną rolę, ponieważ umożliwia zastosowanie religijnych pojęć w różnych okolicznościach – zarówno w zaawansowanym myśleniu teologicznym, jak i w potocznej religijnej interpretacji codziennego doświadczenia.

W myśleniu religijnym występuje także tendencja do tworzenia antropomorficznych reprezentacji. Antropomorfizacja istot nadprzyrodzonych rozumiana jako przypisywanie tym istotom cech ludzkich jest zjawiskiem dawno dostrzeżonym przez filozofów i badaczy religii. Jednak kognitywne teorie religijnej antropomorfizacji rozwinięte zostały całkiem niedawno. Należy do nich teoria zaproponowana przez Stewarta Guthrie'ego (1993), zgodnie z którą antropomorfizacja stanowi naturalną i uniwersalną skłonność umysłu ludzkiego. Skłonność do antropomorfizacji jest prawdopodobnie adaptacją umożliwiającą szybkie, choć nie zawsze trafne wykrywanie zagrożeń i szans w środowisku. Stosowanie modelu antropomorficznego do rozumienia świata dokonuje się niejako automatycznie, w sposób nierefleksyjny. Człowiek posiada intuicyjną zdolność rozumienia innych osób i zdolność tę wykorzystuje w rozumieniu świata pozaludzkiego. Skłonność do antropomorfizacji jest wprawdzie powszechna i nie ogranicza się do myślenia religijnego, jednak antropomorficzne przekonania religijne są bardziej usystematyzowane i zintegrowane, a niekiedy również usankcjonowane w ramach doktryny.

Badacze kognitywni argumentują, że pojęcia religijne nie są tworzone w sposób przypadkowy, lecz zgodnie z uniwersalnym schematem, określonym przez skłonności i ograniczenia poznawcze ludzkiego umysłu (Boyer, 1994; Boyer, 2005). Powierzchnowy przegląd olbrzymiego zasobu idei religijnych wytworzonych w tysiącach znanych religii świata sprawia wrażenie, że wyobraźnia religijna nie ma granic. Wrażenie to jest jednak błędne, co wykazuje dogłębna analiza struktury przedstawień religijnych. Idee istot nadprzyrodzonych tworzone są za pomocą ograniczonego katalogu cech. Istotną rolę w tym procesie odgrywają schematy pojęciowe, za pomocą których ludzkie umysły kategoryzują świat, opierając się na intuicyjnych oczekiwaniach dotyczących tego, jakie rzeczy znajdują się w świecie. Taka intuicyjna wiedza jest zdaniem wielu badaczy w znacznej mierze niezależna od kultury. W rezultacie pojęcia religijne, choć nie są kulturowo uniwersalne, powstają na bazie

intuicyjnego rozumienia świata, które jest wytworem uniwersalnych procesów poznawczych (Boyer, 1994, s. viii).

Chociaż pojęcia istot nadprzyrodzonych bazują na intuicyjnej wiedzy, zawierają także cechy kontrintuicyjne, co oznacza, że naruszają przynajmniej niektóre intuicyjne oczekiwania dotyczące tego, jak się mają rzeczy w świecie (Boyer, 2005). Kognitywiści wskazują, że nasze potoczne myślenie o świecie bazuje na tzw. intuicyjnej ontologii. Jest to wiedza, która w odróżnieniu od wiedzy naukowej nie jest wytworem refleksji, lecz powstaje w wyniku działania automatycznych mechanizmów wnioskowania. Zwykle nie uświadamiamy sobie faktu, że nasze codzienne zachowania oparte są na „intuicjach”, które pozwalają nam sprawnie działać bez zastanowienia: opuszczając pokój, kierujemy się w stronę drzwi, a nie w stronę ściany, ponieważ nasza intuicyjna fizyka podpowiada nam, że ciało stałe nie przenika przez inne ciało stałe. Również idee religijne powstają na bazie intuicyjnej wiedzy o świecie – potocznej fizyki, biologii czy psychologii – mają jednak również takie cechy, które przeczą intuicyjnym wyobrażeniom. Duch na przykład jest osobą, która podobnie jak inne osoby posiada stany umysłowe – postrzeżenia, przekonania, pragnienia i intencje – i w tym względzie przystaje do intuicyjnej psychologii leżącej u podłoża naszego myślenia o osobach. Posiada jednak również takie cechy, które przeczą intuicyjnej wiedzy, na przykład zdolność przenikania przez ściany. Cechy nieprzystające do intuicyjnej wiedzy sprawiają, że pojęcie ducha przyciąga uwagę, natomiast cechy zgodne z intuicyjną ontologią pobudzają umysłowe systemy wnioskowania.

Badania prowadzone nad zapamiętywaniem pojęć wskazują, że pojęcia zawierające pojedyncze cechy niezgodne z wiedzą intuicyjną są lepiej zapamiętywane niż pojęcia intuicyjne oraz takie pojęcia, które zawierają dużą liczbę cech niezgodnych z wiedzą intuicyjną. Ponieważ lepiej zapamiętywane idee mają większą szansę na rozpowszechnienie w społeczności ludzkiej, istnieje prawdopodobieństwo, że uzyskują one przewagę nad gorzej zapamiętywanymi ideami (Atran, 2013). Zdaniem badaczy zajmujących się psychologicznymi determinantami kultury funkcjonowanie umysłu ma istotny wpływ na przekaz kulturowy. W szczególności ograniczenia poznawcze, w tym ograniczenia pamięci, sprawiają, że niektóre typy reprezentacji umysłowych uzyskują przewagę w przekazie kulturowym, a przez to łatwiej ulegają rozpowszechnieniu. Idee religijne należą właśnie do takich reprezentacji.

Należy jednak pamiętać o tym, że kognitywne modele przekazu kulturowego zazwyczaj pomijają pozapoznawcze determinanty przekazu, takie na przykład jak technologia pisma, której rozpowszechnienie istotnie zmienia sposób, w jaki idee religijne są rozpowszechniane. Pismo oraz różne artefakty materialne stosowane w religiach stanowią zewnętrzne wobec organizmu rezerwuary pamięci, istotnie poszerzające pamięć biologiczną. Religie kultur piśmiennych umożliwiają tworzenie i rozpowszechnianie wysoce kontrintuicyjnych idei religijnych oraz rozbudowanych systemów teologicznych. Pojęcia Boga tworzone w ramach zrjonalizowanych systemów teologicznych cechuje wysoka kontrintuicyjność, co oznacza znaczną liczbę cech niezgodnych z intuicją zawartych w takich pojęciach Boga.

Pojęcia i przekonania religijne stanowią ważny, ale nie jedyny wymiar religii. Istotną rolę odgrywają w religiach zachowania rytualne, tj. zachowania symboliczne poddane ściśle określonym regułom, powtarzalne, udratyzowane, niezorientowane na cele praktyczne i zwykle mające odniesienie do wierzeń w nadprzyrodzone istoty i moce. Aby odpowiedzieć na pytanie, dlaczego ludzie podejmują rozmaite działania rytualne i jakie funkcje rytuały pełnią w ludzkich społecznościach, rozwinięto w badaniach nad rytuałem różne podejścia: funkcjonalne, symboliczne, fenomenologiczne, strukturalne czy psychoanalityczne. Zwracano uwagę m.in. na komunikacyjny charakter zachowań rytualnych, wskazując, że rytuały służą przekazywaniu informacji. W związku z tym wiele uwagi poświęcano interpretacji rytuałów, mniej zaś wyjaśnieniu działań rytualnych, w szczególności odpowiedzi na pytania: Dlaczego ludzie podejmują działania rytualne? Dlaczego jedne rytuały są często powtarzane (np. Eucharystia), inne zaś wykonywane są tylko jednokrotnie (np. chrzest)? Dlaczego skutki niektórych działań rytualnych są trwałe, innych zaś zanikają z upływem czasu? Dlaczego wreszcie niektóre rytuały wywołują silne reakcje emocjonalne, natomiast inne są monotonne, by nie powiedzieć nudne? Kognitywne teorie rytuału poszukują na te pytania odpowiedzi, formułując sprawdzalne empirycznie hipotezy, które następnie są weryfikowane na podstawie obserwacji rytuałów występujących w różnych tradycjach religijnych.

Działania rytualne różnią się od podejmowanych w życiu codziennym działań instrumentalnych, w których praktycznie określony cel osiągnąć jest poprzez dobór odpowiednich środków działania. Aby otworzyć drzwi

mojego domu, muszę wykonać szereg racjonalnych działań: odnaleźć pęk kuczy, wybrać z niego klucz pasujący do zamka, wykonać odpowiednie ruchy, nacisnąć klamkę, popchnąć drzwi itp. W rytuale tego typu racjonalność działania schodzi na dalszy plan; działania i ich rezultaty nie są powiązane przyczynowo, a uczestnicy rytuału zwykle nie potrafią wyjaśnić, w jaki sposób podjęte działania doprowadziły do założonych rezultatów, lub dają niejasne i tajemnicze odpowiedzi. Wiąże się z tym jeszcze inna ważna cecha odróżniająca rytuały od podejmowanych w życiu codziennym działań: jeśli działanie rytualne nie osiąga zamierzonego celu, nie jest to powodem do jego zakwestionowania i zaniechania w przyszłości. Rytuał mający sprowadzić deszcz z dużym prawdopodobieństwem będzie powtarzany w przyszłości, nawet jeśli wielokrotnie w przeszłości nie wywołał zamierzonego skutku. W życiu codziennym działania podlegają innym regułom: jeśli klucz nie pasuje do zamka, dobieram inny klucz, zaś niepasujący odkładam na bok jako nieprzydatny. W konsekwencji rytuał jest niewielkim stopniu podatny na zmianę.

Robert N. McCauley i Thomas Lawson (2002) zaproponowali kognitywną teorię działań rytualnych, która wyjaśnia rytuał w kategoriach reprezentacji umysłowych. Teoria ta postuluje istnienie swoistej kompetencji rytualnej, na podobieństwo kompetencji językowej badanej m.in. przez Noama Chomsky'ego. Działania rytualne są reprezentowane w umysłach na podstawie systemu, który umożliwia reprezentację zwyczajnych, codziennych działań. Innymi słowy ludzie intuicyjnie rozumieją rytuał jako szczególny typ działania, posiadający określoną formę. Na formę działania składają się trzy elementy: podmiot działania, samo działanie oraz przedmiot działania. Rozumienie rytuału religijnego opiera się na naszych intuicjach dotyczących działania w ogóle, ale obejmuje też specyficzne elementy, takie jak pojęcie nadprzyrodzonej sprawczości oraz ponadnaturalnych skutków działania. W rytuałach religijnych używane są też rozmaite narzędzia, których umieszczenie w strukturze działania wpływa na rozumienie rytuałów przez uczestniczące w nich osoby. Rytuał posiada zatem następującą strukturę: osoba A działa na osobę lub przedmiot B za pośrednictwem narzędzia N w celu spowodowania pewnych nadprzyrodzonych skutków dzięki nadprzyrodzonej sprawczości. Obecność kulturowo postulowanej sprawczości nadprzyrodzonej odróżnia rytuał religijny od innych działań. Rytuały różnią się tym, jakiemu elementowi w strukturze rytuału

została przypisana nadprzyrodzona sprawczość. Stosując to kryterium, wyróżnić można dwa podstawowe typy rytuału: taki, w którym nadprzyrodzona sprawczość związana jest podmiotem działania, oraz taki, w którym sprawczość ta związana jest z przedmiotem lub narzędziem działania rytualnego. Informacja o umiejscowieniu nadprzyrodzonej sprawczości w strukturze rytuału jest istotna, ponieważ pozwala przewidywać cechy rytuału takie jak jego powtarzalność oraz częstość powtarzania, poziom pobudzenia emocjonalnego uczestników, trwałość rezultatów działań rytualnych, a także umiejscowienie rytuału (centralne bądź peryferyjne) w danej tradycji religijnej. Przykładowo McCauley i Lawson wskazują, że rytuały, w których sprawczość przypisana jest podmiotowi działania, są rzadziej powtarzane, ich skutki uważane są za trwałe, wywołują u uczestników silne reakcje emocjonalne, a także postrzegają się je jako kluczowe w obrębie danej tradycji. Przykładem takiego działania może być chrzest w tradycji chrześcijańskiej, w trakcie którego podmiot działania (kapłan) udziela sakramentu. Rytuał ten nie wymaga wielokrotnego powtarzania, jego skutki zachowują trwałość i należy on do kluczowych obrzędów chrześcijańskich. Odwrotnie rytuały, w których nadprzyrodzona sprawczość powiązana jest z przedmiotem lub narzędziem działania – są one częściej powtarzane, a ich rezultaty są mniej trwałe, nie wywołują silnych emocji i zwykle nie odgrywają kluczowej roli. Przykładem takiego rytuału może być składanie ofiary bogu.

Rytuały można także różnicować ze względu na przejawiające się w nich style religijności opisane przez brytyjskiego antropologa Harveya Whitehouse'a. Wyróżnił on dwa podstawowe style religijności – obrazowy i doktrynalny – które wprawdzie mogą współistnieć w poszczególnych tradycjach religijnych, ale bazują na odmiennych mechanizmach poznawczych, w szczególności na odmiennych rodzajach pamięci (Whitehouse, 2004). Whitehouse poświęcił szczególną uwagę temu, jak poszczególne style religijności wpływają na kulturowy przekaz idei i zachowań religijnych. Obrazowy styl religijności obejmuje rytuały, które powtarzane są rzadko, ale cechują się bardzo dużą widowiskowością. Wywołują silne wrażenia w uczestnikach i są zapamiętywane dzięki pamięci epizodycznej. Zwiększają też spójność społeczną i sprawiają, że jednostka silnie utożsamia się z grupą społeczną. Religie, w których dominuje styl obrazowy, nie wykształcają silnego przywództwa religijnego, hierarchii oraz ortodoksji doktrynalnej. Ponieważ przywiązane są

do lokalnego kontekstu kulturowego, z trudem rozpowszechniają się poza tym kontekstem. Przykładem rytuału typowego dla tego typu religijności jest rytuał inicjacji praktykowany m.in. w kulturach tradycyjnych. W trakcie takiego rytuału inicjowany poddawany jest działaniom, które zwykle pamięta przez całe życie. Jednocześnie rytuał inicjacji, w wyniku którego chłopiec staje się mężczyzną, jest rytuałem jednorazowym, niepowtarzalnym. Z kolei doktrynalny styl religijności cechuje częste powtarzanie rytuałów oraz przekazu doktrynalnego. O ile przekaz doktrynalny opiera się na pamięci semantycznej, powtarzanie rytuałów bazuje na niejawniej pamięci proceduralnej, co sprawia, że rytuały rzadko stają się przedmiotem refleksji i rzadko podlegają innowacji. Sprzyja to utrzymaniu ortodoksji, która dodatkowo podtrzymywana jest przez ekspertów religijnych, dostarczających autorytatywnej interpretacji doktryny i rytuału. Często powtarzanym rytuałom – podobnie jak w koncepcji Lawsona i McCauleya – nie towarzyszą w tym stylu religijności intensywne emocje. Ponieważ treści religijne przekazywane są w formie doktryny, czasami przy użyciu pisma, religie oparte na tym stylu religijności łatwiej rozprzestrzeniają się w różnych środowiskach kulturowych. Zorganizowane kościoły chrześcijańskie byłyby dobrym przykładem instytucji opartych na doktrynalnym stylu religijności. Należy jednak podkreślić, że dwa style religijności – obrazowy i doktrynalny – często współwystępują w ramach poszczególnych tradycji religijnych, zaś dominacja jednego z nich pozwala określić tradycję jako obrazową lub doktrynalną.

Procesy poznawcze leżące u podłoża religii odgrywają kluczową rolę w kształtowaniu religijnego zachowania. Jednak badanie procesów poznawczych dotyczy nie tylko tego, co dzieje się w naszych głowach. Niektórzy badacze kognitywni argumentują, że przetwarzanie poznawcze obejmuje także elementy środowiska znajdującego się na zewnątrz organizmu. Traktują oni poznanie ludzkie jako proces, który przebiega nie tylko w mózgu, ale ponadto w ciele oraz w świecie otaczającym organizm. Człowiek wykształcił zdolność posługiwania się artefaktami materialnymi jako narzędziami myślenia. Nie ma zgody co do tego, czy owe zewnętrzne wobec organizmu artefakty należy ujmować jako elementy systemu poznawczego, czy raczej należy mówić o interakcji systemu poznawczego ze środowiskiem. Nie ulega jednak wątpliwości, że kultury ludzkie w szerokim zakresie wykorzystują zdolność delegowania pewnych zadań poznawczych na zewnątrz oraz przetwarzania

informacji na podstawie różnych elementów wytworzonego przez człowieka środowiska. W tym kontekście myślenie religijne może być rozumiane jako proces obejmujący zarówno działanie systemu poznawczego, jak i interakcje systemu poznawczego ze środowiskiem.

Przedmioty materialne – zarówno te będące częścią świata natury, jak i te wytworzone przez człowieka – stanowią nieodłączny wymiar każdej niemal religii. Do tych pierwszych należą sakralizowane obiekty natury – święte góry, rzeki, drzewa, kamienie, ciała niebieskie, zwierzęta, do tych drugich – materialne artefakty takie jak obrazy, ikony, relikwie, szaty liturgiczne, świątynie itp. Sakralizacji może ulec w zasadzie każdy przedmiot materialny. I rzeczywiście, w dziejach religii świata rozmaite przedmioty uznawane były za święte. Materialne artefakty nie są tylko czymś w rodzaju otoczki czy dekoracji życia religijnego, ale pełnią szereg istotnych funkcji poznawczych (Sztajer, 2018). Można je potraktować jako narzędzia, które wspierają, a czasami umożliwiają tworzenie, rozumienie, przetwarzanie i komunikowanie idei religijnych. Materialne reprezentacje istot nadprzyrodzonych niejako uobecniają niedostępne zmysłom nadprzyrodzone światy, przedstawiając w konkretnej, namacalnej postaci to, co wcześniej było przedmiotem nieokreślonych wyobrażeń.

Jakie funkcje poznawcze pełnią przedmioty materialne w religii? Po pierwsze, umożliwiają uzewnętrznienie z natury subiektywnej świadomości religijnej i wprowadzenie treści tej świadomości do publicznego obiegu. Jest to niezwykle ważna funkcja, ponieważ religie są przedsięwzięciami zbiorowymi, podzielanymi przez członków danej wspólnoty. Przedstawianie wyobrażeń religijnych w formie materialnej sprawia, że postrzegane są one jako należące do rzeczywistości zewnętrznej wobec podmiotu, już nie tylko subiektywne, lecz intersubiektywne. Po drugie, uzewnętrznienie idei religijnych wiąże się z ich zakotwiczeniem w bardziej stabilnych reprezentacjach materialnych. Powiązanie struktur pojęciowych z odpowiadającymi im strukturami materialnymi nadaje tym pierwszym większą trwałość. Wyznawcy mogą posługiwać się różnymi wyobrażeniami bóstwa, ale odniesienie tych wyobrażeń do zewnętrznej wobec umysłu reprezentacji materialnej, np. rzeźby bóstwa, nadaje tym wyobrażeniom większą trwałość, zwłaszcza gdy pod uwagę brana jest wielopokoleniowa ciągłość tradycji religijnej. Po trzecie, materialne artefakty poszerzają możliwości pamięci biologicznej. Religie, podobnie

jak inne dziedziny kultury, bazują na zewnętrznych wobec organizmu rezerwuarach pamięci, poszerzających zarówno pamięć indywidualną, jak i zbiorową. Po czwarte, materialne artefakty umożliwiają nadawanie konkretnej postaci abstrakcyjnym ideom religijnym, niejako kondensując zmysłowo nieuchwytny świat nadprzyrodzony do postaci namacalnej, dostępnej zmysłom. Dzięki temu przetwarzanie reprezentacji religijnych obejmuje procesy sensomotoryczne. Po piąte, materialne artefakty sprawiają, że świat nadprzyrodzony jawi się jako dostępny, obecny tu i teraz, a przede wszystkim bardziej rzeczywisty. Poczucie rzeczywistości nadprzyrodzonych istot i mocy jest indukowane poprzez powiązanie idei religijnych z uchwytnymi praktycznie przedmiotami, które nie tylko są dostępne zmysłom, ale którymi można również manipulować w rytuale. Niektórzy badacze wskazują ponadto, że materialne artefakty, zwłaszcza fizyczne reprezentacje istot nadprzyrodzonych, we wczesnych etapach rozwoju religii mogły funkcjonować nie tyle jako odzwierciedlenie pojęć religijnych, ile jako ich pierwowzory. Innymi słowy idee religijne mogły powstawać jako reprezentacje przedmiotów materialnych, niejako wyłaniać się z kontaktu człowieka z takimi przedmiotami.

Refleksja systematyczna z wnioskami i rekomendacjami

Kognitywne podejście w badaniach nad religiami ulegało w ostatnich dekadach licznym modyfikacjom, które świadczą o ekspansywności tego programu badawczego. Podstawową zaletą tego podejścia jest jego interdyscyplinarność oraz dążenie do uwzględnienia różnych warunkowań, jakim podlegają przekonania i zachowania religijne. Obecnie podejście kognitywne zmierza ku integracji wiedzy z zakresu kognitywistyki, neuronauki, teorii ewolucji i nauk o kulturze. Zaletą jest także formułowanie nowych hipotez, które poddawane są systematycznym i powtarzalnym testom empirycznym, w tym testom eksperymentalnym. Zastosowanie metod przyjętych w naukach o poznaniu i ewolucji do badania zjawisk religijnych zbliża badania religioznawcze do przyjętych standardów naukowości.

Krytycy podejścia kognitywnego zwracają uwagę na szereg jego ograniczeń. Formułowane przez nich zarzuty dotyczą przede wszystkim

redukcjonizmu będącego istotnym składnikiem kognitywnego wyjaśniania religii, zastosowania zbyt wąskiego podejścia i ograniczania się do jednego typu czynników, marginalizacji interpretacji i rozumienia na rzecz wyjaśniania zjawisk religijnych w kategoriach mechanizmów przyczynowych (zob. Cohen et al., 2008). Niektóre z tych zarzutów wydają się słuszne w odniesieniu do wczesnych kognitywnych teorii religii, lecz nieuzasadnione, gdy pod uwagę wzięte zostaną nowsze kierunki badań i osiągnięcia religioznawstwa kognitywnego. W nowszych badaniach wielowymiarowość i silne zróżnicowanie zjawisk religijnych, a także kwestie rozumienia i interpretacji nie są ignorowane. Zarzut redukcjonizmu kierowany pod adresem kognitywnych teorii religii jest słuszny w tym sensie, że wyjaśnienie zjawisk religijnych polega na ich redukowaniu do zjawisk innego typu. Twórcy takich teorii odrzucają rozpowszechnione wśród fenomenologów religii założenie o autonomii religii, nieredukowalności zjawisk religijnych. Należy jednak pamiętać, że redukcjonizm występujący w podejściu kognitywnym zasadniczo różni się od redukcjonizmu właściwego niektórym klasycznym teoriom religii, które usiłowały w pełni wyjaśnić religię, opierając się na zjawiskach niższego rzędu – społecznych, psychicznych, ekonomicznych etc. W skrajnej postaci redukcjonizm taki prowadził do kwestionowania nie tylko odrębności i specyfiki, ale także realności religii. Kognitywne teorie religii umożliwiają tworzenie wielowymiarowych modeli zjawisk religijnych, uwzględniających różne typy uwarunkowań. Ważną rolę odgrywa idea pluralizmu eksplanacyjnego (wielości sposobów wyjaśniania), która w odniesieniu do religii oznacza, że nie sposób wyjaśnić w sposób pełny i ostateczny tak złożonego i wielowymiarowego zjawiska, jakim jest religia. Zamiast tego proponuje się uwzględnienie różnych poziomów analizy zjawiska oraz zastosowanie wielu różnych metod.

Kognitywne badania nad religią nie tylko pogłębiają rozumienie zjawisk religijnych, ale także dostarczają wiedzy, która może być wykorzystana w praktyce, zwłaszcza w obszarze edukacji religijnej, działalności grup religijnych czy polityki wyznaniowej. Religioznawstwo kognitywne stanowi ponadto punkt odniesienia w sporach światopoglądowych dotyczących religii, wątpliwe jest jednak, by dostarczana przez nie wiedza umożliwiała rozstrzygnięcie takich sporów.

BIBLIOGRAFIA

- Atran, S. (2013). *Ewolucyjny krajobraz religii*. Kraków: Zakład Wydawniczy „Nomos”.
- Barrett, J.L. (2004). *Why would anyone believe in God?*. Walnut Creek, CA: AltaMira Press.
- Barrett, J.L. (Red.). (2022). *Oxford Handbook of the Cognitive Science of Religion*. Oxford: Oxford University Press.
- Boyer, P. (1994). *The naturalness of religious ideas: A cognitive theory of religion*. Berkeley: University of California Press.
- Boyer, P. (2005). *I człowiek stworzył bogów... Jak powstała religia?* Przeł. K. Szeżyńska-Maćkowiak. Warszawa: Prószyński i S-ka.
- Cohen, E., Lanman, J., McCauley, R., & Whitehouse, H. (2008). Common Criticisms of the Cognitive Science of Religion-Answered. *The CSSR Bulletin*, 37, 112–115.
- Guthrie, S.E. (1993). *Faces in the clouds: A new theory of religion*. New York: Oxford University Press.
- Lawson, E.T., & McCauley, R.N. (1990). *Rethinking religion: Connecting cognition and culture*. Cambridge, UK: Cambridge University Press.
- Liddle, J.R., & Shackelford, T.K. (Red.). (2021). *The Oxford Handbook of Evolutionary Psychology and Religion*. Oxford: Oxford University Press.
- McCauley, R.N., & Lawson, E.T. (2002). *Bringing Ritual to Mind: Psychological Foundations of Cultural Forms*. Cambridge: Cambridge University Press.
- McCauley, R.N. (2011). *Why Religion is Natural and Science is Not*. Oxford: Oxford University Press.
- Schjoedt, U. (2009). The religious brain: A general introduction to the experimental neuroscience of religion. *Method & Theory in the Study of Religion*, 21(3), 310–339.
- Sztajer, S. (2018). *Pojęciowa konstrukcja świata religijnego. Przekonania religijne a procesy poznawcze*. Poznań: Wydawnictwo Naukowe Wydziału Nauk Społecznych UAM.
- White, C. (2021). *An Introduction to the Cognitive Science of Religion: Connecting Evolution, Brain, Cognition, and Culture*. Milton: Taylor and Francis.
- Whitehouse, H. (2004). *Modes of religiosity: A cognitive theory of religious transmission*. Walnut Creek, Calif.; Oxford: Altamira Press.

Michał Wierzchoń
Uniwersytet Jagielloński w Krakowie
<https://orcid.org/0000-0002-7347-2696>
<https://doi.org/10.35765/slowniki.486>

Substytucja sensoryczna

Streszczenie

DEFINICJA POJĘCIA: Tekst omawia zjawisko substytucji sensorycznej. W pierwszej części rozdziału przedstawiono rozważania dotyczące definicji pojęcia, wykazano warunki konieczne dla wystąpienia substytucji oraz poddano analizie przykłady tego zjawiska.

ANALIZA HISTORYCZNA POJĘCIA: Zaprezentowano krótki rys historyczny rozwoju technologii substytucji sensorycznej. Wskazano również potrzebę tworzenia nowych, nieinwazyjnych urządzeń protetycznych konstruowanych z uwzględnieniem potrzeb osób niewidomych.

UJĘCIE PROBLEMOWE POJĘCIA: Dyskusji poddano trzy obszary badań nad substytucją sensoryczną, tj. 1) badania nad rozwojem urządzeń substytucji, 2) badania neurobiologiczne dokumentujące neuroplastyczność będącą efektem wykorzystywania urządzeń substytucji oraz 3) badania nad doświadczeniem subiektywnym substytucji.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Zaproponowano kierunki dalszych badań nad analizowanym zjawiskiem, które powinny przybliżyć możliwość wprowadzenia technologii substytucji sensorycznej do codziennego użytku.

Słowa kluczowe: substytucja sensoryczna, doświadczenie, percepcja, neuroplastyczność, rehabilitacja osób niewidomych

Definicja pojęcia

Substytucja sensoryczna pozwala osobie pozbawionej dostępu do informacji o środowisku osiągalnej w sytuacjach typowych dzięki użyciu określonego zmysłu (np. informacji wzrokowej w przypadku osób niewidomych) dostarczyć ją za pośrednictwem innej modalności zmysłowej (np. słuchu lub dotyku). Przykładem substytucji może być sensor parkowania w samochodzie. W tym przypadku informacja o odległości przeszkody zazwyczaj dostępna wzrokowo (zmysł substytuowany) staje się dostępna dzięki zastosowaniu informacji dźwiękowej dostępnej słuchowo (zmysł substytuujący). Słuch zastępuje zatem modalność wzrokową.

Współcześnie termin „substytucja sensoryczna” stosowany jest w odniesieniu do udostępniania informacji z wykorzystaniem technologii protetycznych (urządzenia substytucji sensorycznej, *sensory substitution devices*). Większość urządzeń substytucji sensorycznej składa się z sensora (np. kamery), stymulatora (np. słuchawek) oraz urządzenia implementującego algorytm translacji sygnału sensora. Dzięki tym urządzeniom sygnał z sensora ma się stać dostępny (i zrozumiały) dla zmysłu substytuującego. Zastosowanie urządzeń substytucji sensorycznej pozwala zatem na prezentację wybranych informacji dostępnych zazwyczaj dzięki zmysłowi substytuowanemu za pomocą zmysłu substytuującego. Jak już wspomniano, najprostszą ilustracją zastosowania substytucji sensorycznej jest sensor parkowania samochodu. W analizowanym przypadku niedostępna wzrokowo informacja na temat obiektów znajdujących się za samochodem zostaje udostępniona dzięki zastosowaniu sensora parkowania, prostego algorytmu, w którym informacja na temat odległości przekładana jest na częstotliwość dźwięków oraz efektora w postaci głośnika emitującego wygenerowane przez algorytm dźwięki.

Rozwój technologii substytucji sensorycznej powiązany jest przede wszystkim z ich funkcją rehabilitacyjną. Urządzenia substytucji sensorycznej pozwalają dostarczyć osobom niewidomym informacje na temat charakterystyk środowiska poznawanych zazwyczaj za pomocą wzroku, np. odległość od obiektu, jego kształt czy też kolor (Kałwak, Reuter, Łukowska, Majchrowicz & Wierzchoń, 2018). Większość badań z tego zakresu opisywanych w literaturze prowadzona była z udziałem osób

niewidomych od urodzenia. Substytucję sensoryczną można jednak obserwować również u osób, które utraciły wzrok w późniejszym etapie życia, a nawet u osób widzących, którym czasowo ograniczono dostęp do informacji zmysłowej (Łukowska i współprac., 2021).

Analiza historyczna pojęcia

Pojęcie substytucji sensorycznej zostało wprowadzone do literatury naukowej przez Paula Bacha-y-Ritę w latach 80. XX wieku. Integruje ono obszar wcześniej prowadzonych badań nad rozwojem zróżnicowanych rozwiązań technologicznych mających na celu wspomaganie percepcji osób niewidomych. Substytucja sensoryczna możliwa jest dzięki tworzeniu i rozwojowi nowych, nieinwazyjnych urządzeń protetycznych opartych na wspomaganym komputerowo przetwarzaniu informacji sensorycznych niedostępnych dla osób niewidomych. W przeciwieństwie do tradycyjnych rozwiązań wspomagających funkcjonowanie osób niewidomych (np. białej laski), substytucja sensoryczna oparta jest więc na rozwoju technologii. Urządzenia substytucji sensorycznej są niezależne od innych rozwiązań technologicznych szerokiego zastosowania (np. mobilnych technologii asystujących dostępnych za pośrednictwem aplikacji na telefony komórkowe). Celem urządzeń substytucji sensorycznej jest udostępnienie osobie niewidomej niedostępnej informacji wzrokowej za pośrednictwem innego, funkcjonującego u niej prawidłowo zmysłu (np. słuchu bądź dotyku).

Pierwszym urządzeniem opisywanym w literaturze światowej jako urządzenie substytucji sensorycznej był elektroftalm – sztuczne oko – opracowane przez polskiego okulistę, Kazimierza Noiszewskiego pod koniec XIX. Urządzenie przekształcało informacje dotyczące bodźców świetlnych na dotykowe i dźwiękowe. Prace nad zmodyfikowanym prototypem urządzenia zaproponowanego przez Noiszewskiego prowadził w latach 60. XX wieku inny polski okulista, Witold Starkiewicz (Farmer, 1978). Elektroftalm Starkiewicza wykorzystywał zestaw trzystu fotokomórek, których pobudzenie uruchamiało wibratory przytwierdzone do czoła niewidomego. Dzięki temu rozwiązaniu technologicznemu możliwe było stworzenie uproszczonego obrazu otoczenia osoby niewidomej. Koncepcja elektroftalmu i wyniki badań prowadzonych przez Starkiewicza

były publikowane w Polsce i za granicą, jednak informacje o tym nie są zazwyczaj podawane we współczesnych publikacjach naukowych dotyczących substytucji sensorycznej. Publikacje te wymieniają najczęściej Brainport jako pierwsze urządzenie substytucji sensorycznej.

Pierwsza wersja urządzenia Brainport została opisana w czasopiśmie „Nature” w 1969 roku (Bach-y-Rita & Kerchel, 2003). Brainport to urządzenie substytucji typu wzrok-dotyk. Przekształca obraz z wbudowanej kamery niskiej rozdzielczości na wzorce łagodnych impulsów elektrycznych. Urządzenie dostarcza informacji o bodźcu wzrokowym, kodując jasność każdego piksela obrazu i przekładając ją na odpowiadającą jej siłę drgania stymulatora umieszczonego na języku. Stymulator ma rozdzielczość analogiczną do kamery, tj. 20 na 20 pikseli. Użytkownicy Brainportu wyczuwają wzory jasności na powierzchni języka i uczą się rozpoznawać kształt, rozmiar lub lokalizację prezentowanych obiektów. Urządzenie pracuje w trybie on-line, pozwala zatem wykrywać również ruch obiektów w otoczeniu. Brainport nie dostarcza natomiast informacji o głębi obrazu (na przykład odległości obiektów). Urządzenie zostało skomercjalizowane i od 1998 jest dostępne na rynku. W 2015 wersja V100 została zaakceptowana do użytku przez Food and Drug Administration (FDA). Urządzenie jest więc dostępne w sprzedaży, choć nie używa się go powszechnie.

Brainport został skonstruowany jako urządzenie wspierające poczucie równowagi, szczególnie u pacjentów po udarach. Z wykorzystaniem tego urządzenia podjęto pionierskie badania dotyczące neuroplastyczności mózgu osób niewidomych od urodzenia. Te prace stały u podstawy wielu eksperymentów z zakresu neuronauki poznawczej. Wykazały one, że w przeciwieństwie do ugruntowanych wcześniej w literaturze neurobiologicznej poglądów organizacja funkcjonalna mózgu nie jest niezmienna, ale może podlegać daleko idącej reorganizacji nawet u osób dorosłych (Heimler & Amedi, 2020).

Ujęcie problemowe pojęcia

Najważniejszym problemem badawczym jest z pewnością rozwój technologii urządzeń substytucji sensorycznej. W realizacji tego zadania konieczne jest podejście interdyscyplinarne, ponieważ analizowana

technologia wymaga wiedzy o zasadach przetwarzania informacji sensorycznej przez mózg, ale również znajomości metod projektowania urządzeń protetycznych, metodologii badań ich użyteczności, kompetencji w zakresie programowania, doświadczenia w badaniach grup osób niewidomych, znajomości metod testowania urządzeń w kontekście interakcji środowiskowej i wielu innych. Badania nad substytucją prowadzą zazwyczaj interdyscyplinarne zespoły badawcze zrzeszające między innymi psychologów, neuronaukowców, informatyków, inżynierów, lekarzy czy też rehabilitantów. Warto zwrócić uwagę, że wypracowanie powszechnie akceptowanego przez użytkowników rozwiązania technologicznego substytucji sensorycznej miałoby fundamentalne znaczenie dla dużej grupy ludzi. Światowa Organizacja Zdrowia szacowała w 2012 roku liczbę osób niewidomych na 40 milionów, podczas gdy znacząca niepełnosprawność zmysłu wzroku dotyczyła 285 milionów ludzi. Wiele z tych osób zamieszkuje w krajach rozwijających się, zatem na powszechną dostępność rozwiązania technologicznego będą mieć wpływ zarówno jego użyteczność, jak i cena. W literaturze opisano wiele rozwiązań technologicznych konkurencyjnych dla opisanego powyżej Brainportu (oraz bardzo do niego podobnego konceptualnie elektroftalmu Noiszewskiego). Przykładami mogą być Enactive Torch, the vOICe czy też Colorophone (Łukowska, Kałwak, Osiński, Janik & Wierzchoń, 2022).

Enactive Torch to urządzenie substytucji typu wzrok-dotyk zaproponowane przez Toma Froese (Froese, McGann, Bigge, Spiers & Seth, 2012). Przekształca ono informację na temat odległości obiektu na drgania. W standardowych ustawieniach zbliżanie się do obiektu skutkuje zwiększoną siłą wibracji, podobnie jak w samochodowych sensorach parkowania. Urządzenie wykorzystuje czujnik podczerwieni lub ultradźwięków. Pracuje również w trybie on-line, dostarczając informacji o odległości od przeszkody. Umożliwia w ten sposób lokalizację obiektu w przestrzeni i oszacowanie jego wielkości. Dzięki temu uczestnicy mogą nawigować w środowisku pomimo braku informacji wzrokowej.

Z kolei the vOICe to urządzenie substytucji typu wzrok-słuch zaproponowane przez Petera Meijera (1992). Przekształca ono obraz z wbudowanej kamery na informację słuchową, dostarczając informacji o kolorze i pozycji obiektów w scenie percepcyjnej. Urządzenie pracuje w trybie off-line, tj. pozwala zaprezentować statyczną scenę percepcyjną

(na przykład zdjęcie lub klatkę stop z obrazu kamery). Scena percepcyjna reprezentowana jest w formie sceny dźwiękowej (*soundscape*). Informacja na temat koloru i pozycji obiektu w scenie reprezentowana jest za pomocą wysokości, głośności i następstwa czasowego dźwięków: wysokość dźwięku pozwala kodować informacje na temat położenia obiektu w wymiarze wertykalnym (wysokie dźwięki oznaczają obiekt położony w górnej części pola wzrokowego), głośność dźwięku pozwala kodować informację na temat jasności (głośny dźwięk reprezentuje bardzo jasny obiekt), następstwo czasowe dźwięków pozwala kodować informację na temat położenia horyzontalnego obiektu (w pierwszej kolejności prezentowane są dźwięki reprezentujące obraz po lewej, a następnie po prawej stronie pola widzenia).

Algorytm kodowania urządzenia The vOICe został rozwinięty przez Amira Amediego w ramach technologii EyeMusic, w której do opisanych powyżej reguł kodowania położenia i jasności obiektu dodano informacje na temat pięciu podstawowych kolorów: białego, niebieskiego, czerwonego, zielonego i żółtego. Poszczególne kolory reprezentowane z wykorzystaniem różnych instrumentów muzycznych. Warto zauważyć, że twórca the vOICe, Peter Meijer, uważa, że mimo pozornego podobieństwa the vOICe i EyeMusic dostosowane są one do innych zadań i stawiają sobie za zadanie odpowiadanie na inne potrzeby osób badanych. Meijer uważa, że podstawową funkcją urządzeń substytucji sensorycznej jest nawigacja w przestrzeni, omijanie przeszkód i rozpoznawanie obiektów. W jego opinii żadne z tych zadań nie wymaga postrzegania kolorów obiektów. Nie jest jasne, na jakiej podstawie twórca the vOICe formułuje swoje tezy. Mają one jednak z pewnością znaczący wpływ na zaproponowane przez niego rozwiązania technologiczne. Obserwacja ta potwierdza, że kluczowe dla rozwoju urządzeń substytucji sensorycznej są nie tylko dobrze działające algorytmy, ale również precyzyjna definicja potrzeb docelowych użytkowników.

Wśród wielu urządzeń substytucji sensorycznej opisywanych w literaturze warto zwrócić uwagę na inne urządzenie typu wzrok-słuch o nazwie Colorophone zaproponowane przez Dominika Osińskiego (Osiński, Łukowska, Bizoń-Angov, Hjelme & Wierzchoń, 2021). Urządzenie to przekłada informację o kolorze obiektów na informację słuchową. Rozpoznawanie kolorów możliwe jest dzięki algorytmowi przetwarzającemu sygnał obrazu kamery na dźwięki prezentowane przez

słuchawki otwarte (tj. niezasłaniające uszu użytkowników). W najnowszej wersji urządzenia, Colorophone 2.2, algorytm wykorzystuje wartości matrycy kolorów RYGB: każdy z czterech kolorów podstawowych (tj. czerwony, żółty, zielony i niebieski) jest przypisany do innej częstotliwości dźwięku (czerwony do wysokiej częstotliwości, żółty i zielony do średnich częstotliwości, a niebieski do najniższej). Zgodnie z regułami matrycy RYGB wszystkie pozostałe kolory są reprezentowane poprzez połączenie kolorów podstawowych. Intensywność każdego komponentu koloru kodowana jest za pomocą głośności dźwięku odpowiadającemu danemu kolorowi podstawowemu. Badani słyszą zatem kilka dźwięków naraz, a ich zadaniem jest określenie konkretnej barwy obiektu. W wersji Colorophone 2.2 urządzenie może działać w dwóch trybach: punktowym bądź panoramicznym (por. Łukowska, Kałwak, Osiński, Janik & Wierzchoń, 2022). W trybie punktowym informacja na temat koloru uśredniana jest jedynie ze zbioru pikseli znajdujących się w centrum pola widzenia (np. 50 × 50 pikseli obrazu). W opisywanym trybie skanowanie całej sceny percepcyjnej możliwe jest zatem tylko poprzez ruchy głową zmieniające obraz w centrum pola kamery. W trybie panoramicznym informacja na temat koloru kodowana jest z przebiegającego przez całe pole widzenia (od lewej do prawej strony) paska pikseli umieszczonego w połowie wysokości sceny (np. 50 pikseli wysokości). Informacje prezentowane w lewym polu widzenia przekładane są na dźwięk w lewej słuchawce, informacje w prawym polu widzenia na dźwięki w prawej słuchawce. Obiekty prezentowane centralnie generują dźwięk stereofoniczny. W trybie panoramicznym możliwe jest zatem uzyskanie informacji na temat układu obiektów w przestrzeni bez konieczności skanowania sceny percepcyjnej. Wykorzystując urządzenie Colorophone w Laboratorium Badania Świadomości Uniwersytetu Jagiellońskiego, przeprowadzono w ostatnich latach szeroko zakrojone badania poświęcone neuronalnym i poznawczym mechanizmom kształtowania doświadczenia substytucji sensorycznej (Łukowska i współprac., 2021; Łukowska, Kałwak, Osiński, Janik & Wierzchoń, 2022).

Badania mające na celu rozwój urządzeń substytucji sensorycznej można zaliczyć do dwóch przenikających się wzajemnie grup: badań behawioralnych lub badań użyteczności. Badania behawioralne mają na celu wykazanie, jakiego typu zadania mogą być skutecznie wykonywane z wykorzystaniem urządzeń substytucji sensorycznej. Koncentrują

się one najczęściej na pomiarze poziomu wykonania zadań. Najczęściej wykorzystywane zadania to rozpoznanie obiektów (np. kształtów geometrycznych, liter, twarzy), lokalizacja obiektów, szacowanie liczby obiektów, nawigacja w przestrzeni, omijanie przeszkód, czytanie czy też rozpoznawanie kolorów (por. np. Ptito i współprac., 2005, Froese i współprac., 2012). Badaczy interesuje poprawność i czas wykonania zadań, które wskazują, na ile wykorzystanie danego urządzenia może przekładać się na poprawę funkcjonowania osób pozbawionych dostępu do danego zmysłu w środowisku naturalnym. Warto zwrócić uwagę, że zadania behawioralne starają się zazwyczaj imitować czynności, do których realizacji urządzenie rzeczywiście mogłoby się przydać osobom niewidomym w życiu codziennym. Co ciekawe, badacze rzadko opracowują zestawy zadań behawioralnych w porozumieniu z osobami niewidomymi, pytając o ich rzeczywiste potrzeby. Bazują zazwyczaj na swoich ukrytych założeniach dotyczących potrzeb osób niewidomych i wynikających z nich wyobrażeniach dotyczących zadań, do których urządzenia substytucji sensorycznej mogłyby się przydać. Kwestia ta wydaje się niezwykle istotna, ponieważ osoby niewidome często dysponują skutecznymi adaptacjami, czyli strategiami radzenia sobie z zadaniami występującymi w środowisku naturalnym. Urządzenia substytucji nie powinny próbować rozwiązywać zadań, dla których istnieją skuteczne adaptacje, ale raczej koncentrować się na tych zadaniach behawioralnych, dla których nie wypracowano jeszcze skutecznych dostosowań, lub próbować łączyć wiele adaptacji w ramach jednego rozwiązania technologicznego.

Skonstruowanie przydatnego dla użytkownika końcowego urządzenia substytucji sensorycznej nie jest również możliwe bez badań użyteczności. Ich celem jest stwierdzenie, czy potencjalni użytkownicy testowanych rozwiązań protetycznych będą chcieli ich używać. W szczególności badaczy interesuje efektywność i sprawność użytkowania urządzeń, ale również satysfakcja z ich użytkowania. Warto podkreślić ten ostatni wymiar oceny. Nawet najbardziej zaawansowane technologicznie rozwiązanie nie zostanie wprowadzone do użytku, jeżeli osoby niewidome nie będą czerpać satysfakcji z ich użytkowania (np. ze względu na nieintuicyjny interfejs użytkownika, nieprzyjemne dźwięki substytuujące obraz albo nieakceptowalny zdaniem użytkowników wygląd urządzenia). W badaniach użyteczności analizuje się, jakie

funkcje urządzeń substytucji sensorycznej powinny zostać rozwinięte, aby użytkownicy końcowi częściej po nie sięgali (Osiński, Łukowska, Bizoń-Angov, Hjelme & Wierzchoń, 2021). Proponuje się również rekomendacje dotyczące na przykład zasad prezentacji urządzeń substytucji sensorycznej użytkownikom końcowym, instrukcji obsługi czy też form treningu użytkowania urządzenia pozwalających na maksymalne wykorzystanie jego możliwości. Badania te są niezwykle istotne, bowiem analizy rynkowe wskazują, że mimo dostępności wielu rozwiązań technologicznych osoby niewidome bardzo rzadko wykorzystują urządzenia substytucji sensorycznej w życiu codziennym. Wydaje się, że może to być związane z koncentracją badaczy na testowaniu samych rozwiązań technologicznych, a nie testowaniu ich użyteczności w warunkach naturalnych w odniesieniu do potrzeb użytkowników końcowych czy też wcześniej wytworzonych przez osoby niewidome adaptacji do niepełnosprawności. Zwracaliśmy na te zależności uwagę w naszych badaniach, proponując model interaktywnego systemu substytucji sensorycznej (Łukowska, Kałwak, Osiński, Janik & Wierzchoń, 2022). W proponowanym ujęciu zwracamy uwagę, iż system substytucji sensorycznej nie składa się jedynie z urządzenia substytucji, ale uwzględniać powinien również charakterystyki środowiska, w którym jest używany, oraz indywidualne oczekiwania i kompetencje użytkownika.

Z perspektywy kognitywistycznej równie interesujące, co samo projektowanie urządzeń substytucji sensorycznej jest określenie mechanizmu neuronalnego odpowiedzialnego za występowanie opisywanego efektu. Badania poświęcone tej tematyce dotyczyły dotychczas przede wszystkim neuroplastyczności mózgu osób poddanych substytucji. Badania neuroobrazowe skupiały się na śledzeniu aktywacji neuronalnej w korze sensorycznej modalności substytuowanej (np. korze wzrokowej) lub substytuującej (np. korze słuchowej lub somatosensorycznej) w trakcie substytucji oraz badaniu zmian aktywacji neuronalnej w tych obszarach po treningu z wykorzystaniem wybranych urządzeń substytucji sensorycznej (Poirier, De Volder & Scheiber, 2007). Aktywacja obszarów zaangażowanych w sytuacjach typowych w przetwarzanie bodźców modalności substytuowanej została udokumentowana w badaniach fMRI i PET zarówno u osób niewidomych, jak i widzących. Udokumentowano na przykład aktywację kory wzrokowej w trakcie prezentacji dźwięków powiązanych w algorytmie substytucji z informacją

wizualną. Zaobserwowano również wzrost łączliwości pomiędzy rejonami odpowiedzialnymi za przetwarzanie informacji substytuującej i substytuowanej. Wyniki badań neurobiologicznych interpretowane są jako dowód, iż substytucja sensoryczna może prowadzić do neuroplastyczności mózgu. Badania dostarczają jednak różnych rezultatów w zależności od grupy osób badanych, na przykład zaobserwowano różne wyniki dla osób, które utraciły wzrok wcześniej i późno w toku życia (Kałwak i współprac., 2018; Poirier, De Volder & Scheiber, 2007). W badaniach naszego zespołu uzyskano z kolei wyniki wskazujące na znaczenie integracji sensomotorycznej przejawiającej się zwiększonym zaangażowaniem sieci grzbietowej w sytuacji skutecznego posługiwania się urządzeniami substytucji sensorycznej u osób widzących pozbawionych tymczasowo dostępu do informacji wzrokowej po długotrwałym treningu substytucji z wykorzystaniem urządzenia Colorophone (Łukowska i współprac., 2021). Wyniki te wskazują, że urządzenie było wykorzystywane przez osoby badane w działaniu, ale neuroplastyczność nie dotyczyła aktywacji rejonów powiązanych z doświadczeniem wzrokowym.

Choć wyniki niektórych badań neuroobrazowych sugerują, że trening (szczególnie długotrwały – tzw. superużytkownicy, *super-users*) wykorzystujący różne typy urządzeń substytucji sensorycznej substytuujących wzrok prowadzi do aktywacji obszarów mózgu zwykle powiązanych z widzeniem (kora wzrokowa), trudno na podstawie tych danych wyciągać wnioski na temat natury doświadczenia percepcyjnego (tj. czy użytkownicy urządzeń substytucji sensorycznej rzeczywiście widzą za pomocą innego zmysłu). Biorąc pod uwagę niespójne wyniki uzyskiwane w badaniach, kluczowa wydaje się ich interpretacja w ramach jednego modelu teoretycznego. Poirier, De Volder i Scheiber (2007) zaproponowali taki model, wyróżniając dwa różne procesy, które mogą odpowiadać za aktywację kory wzrokowej obserwowaną w badaniach substytucji. Pierwszym z procesów jest integracja multimodalna, w której przypadku dzięki neuroplastyczności dochodzi do powiązania aktywacji rejonów wzrokowych z aktywacją wywołaną przez urządzenia substytucji sensorycznej. Taki mechanizm dominuje prawdopodobnie u osób, które wcześniej utraciły wzrok. U osób widzących, używających systemów substytucji w celach eksperymentalnych, aktywuje ona natomiast prawdopodobnie procesy wyobraźni wzrokowej. W proponowanym

ujęciu substytucja może zatem, w zależności od grupy osób badanych, mieć charakter doświadczenia wzrokowego, doświadczenia multimodalnego lub wyobrażenia.

Problematyka substytucji sensorycznej i towarzyszących jej zmianom reprezentacji mentalnych była również podejmowana przez filozofów. Podejmowali oni często niezwykle interesujący wątek charakterystyki subiektywnego doświadczenia substytucji. Koncentrując się na doświadczeniu substytucji, pytali: 1) czego tak naprawdę doświadczają osoby badane w warunkach substytucji? Czy doświadczenie substytucji zmysłu wzroku za pomocą informacji słuchowej przypomina wzrok, słuch, czy jeszcze inne doświadczenie zmysłowe? Literatura przedmiotu potwierdza, że stosunkowo niewiele wiadomo o charakterze doświadczenia substytucji (Auvray & Myin, 2009; Proulx, 2010). Dyskusje nad tym zagadnieniem koncentrują się na: 1) jakościowej charakterystyce subiektywnego doświadczenia i percepcji z użyciem urządzeń substytucji sensorycznej (Block, 2003; Proulx, 2010), 2) inspirowanej podejściem enaktywnym analizie wpływu interakcji ze środowiskiem na doświadczenie (Kałwak i współprac., 2018) oraz 3) rozwoju metod pomiaru doświadczenia używanych do opisu subiektywnych charakterystyk substytucji (Froese i współprac., 2012). Co ciekawe, większość tych analiz prowadzona jest jedynie w formie dyskursu filozoficznego. Badania poświęcone fenomenologii doświadczenia substytucji są niezwykle rzadkie i dotyczą zazwyczaj arbitralnie i przypadkowo dobranych wymiarów doświadczenia.

Mimo braku systematycznych analiz warto zwrócić uwagę na nieliczne, pionierskie badania poświęcone tej tematyce (por. np. Froese i współprac., 2012). Badania tego typu prowadzone były również w Laboratorium Badań Świadomości Uniwersytetu Jagiellońskiego. W badaniach tych koncentrowaliśmy się na analizach warunków brzegowych doświadczenia substytucji. Sprawdzaliśmy, czy możliwe jest wywołanie u osób widzących doświadczeń wykraczających poza charakterystyki zmysłu substytuującego, to jest na przykład czy możliwe jest „widzenie” za pomocą dotyku. Interesowała nas również możliwość występowania doświadczeń wykraczających poza zaangażowane modalności zmysłowe (por. np. badania postulujące, iż substytucja sensoryczna może prowadzić do powstawania syntetycznych doświadczeń synestetycznych – por. np. Proulx, 2010). Wyniki badań przeprowadzonych

w grupie osób widzących (Reuter, Kałwak, Sznajder & Wierzchoń, w recenzji) pozwoliły nam zaobserwować, że już krótkotrwały trening z urządzeniem Enactive Torch pozwala na skuteczne wykorzystanie go do wspierania wykonania konkretnych działań w toku interakcji środowiskowej. Nie udało się nam jednak zaobserwować systematycznych zmian w zakresie jakościowych charakterystyk raportowanego doświadczenia. Trening nie prowadził zatem do wykształcenia się jakościowo nowej formy doświadczenia substytucji. Raporty badanych dotyczyły najczęściej różnych strategii pozwalających na wykorzystywanie urządzenia do wykonywania stawianych przed nimi zadań (np. odnalezienia przedmiotu). Jedyne wyniki wskazujące na odmienne charakterystyki doświadczenia po treningu z urządzeniami substytucji udało nam się uzyskać w analizach jakościowych, gdzie wykorzystywane do opisu doświadczenia analogie wskazywały, że niektóre osoby badane zaczynały doświadczać sensorycznie zbliżania się do obiektu (np. metafora zbliżania się do pieca). Nasze obserwacje, a także analiza literatury wskazują, że wykształcenie odrębnego jakościowo doświadczenia substytucji może wymagać długotrwałego, systematycznego, ale również dostosowanego do indywidualnych preferencji użytkownika treningu.

Refleksje systematyczne z wnioskami i rekomendacjami

Podsumowując, dotychczasowe badania nad substytucją sensoryczną wykazały, że badani mogą wykorzystywać urządzenia substytucji sensorycznej do wykonywania prostych zadań behawioralnych w środowiskach eksperymentalnym oraz naturalnym. Badania neuroobrazowe wskazują jednocześnie, że substytucja sensoryczna stanowi interesujący model eksperymentalny badania różnych aspektów neuroplastyczności. Dzięki substytucji możliwe jest wywołanie zmiany funkcjonalnej organizacji kory oraz przełamanie silnej specjalizacji poszczególnych rejonów kory poprzez wzrost integracji multimodalnej. Wyniki badań wskazują zatem na potencjalną przydatność technologii substytucji sensorycznej w rehabilitacji osób niewidomych, jednak wykorzystanie urządzeń w praktyce nie jest szeroko rozpowszechnione. Wydaje się, że badania nad substytucją sensoryczną powinny w przyszłości

skoncentrować się na rozwoju badań użyteczności urządzeń substytucji sensorycznej oraz na próbie odpowiedzi na pytanie o jakościową charakterystykę doświadczenia substytucji. Badania użyteczności powinny brać pod uwagę zindywidualizowane ujęcie potrzeb użytkowników docelowych urządzeń substytucji sensorycznej, tak aby konstruowane urządzenia odpowiadały na rzeczywiste potrzeby niewidomych. Aby było to możliwe, badania użyteczności powinny koncentrować się nie tylko na urządzeniu substytucji, ale całym systemie substytucji, to jest również charakterystyce użytkownika oraz środowiska, w którym urządzenie substytucji sensorycznej będą używane. Tylko dzięki analizie całego systemu możliwe będzie zaproponowanie urządzeń i ich funkcji odpowiadających na rzeczywiste potrzeby osób niewidomych. Tego typu działania wydają się konieczne, zważywszy, że pierwsze urządzenie substytucji sensorycznej zaproponowano niemal pół wieku temu, a wciąż nie są one regularnie używane przez osoby niewidome (Łukowska, Kałwak, Osiński, Janik & Wierzchoń, 2022). Znaczenie opisu subiektywnego, indywidualnego doświadczenia substytucji oraz możliwych różnic indywidualnych tego doświadczenia jest mniej oczywiste, ale wydaje się równie kluczowe dla rozwoju tego obszaru badań. Badania w tym obszarze poza próbą charakterystyki własności fenomenalnych doświadczenia substytucji powinny podjąć temat przydatności tego doświadczenia dla interakcji środowiskowej i codziennego funkcjonowania jednostki. Warto dowiedzieć się nie tylko, jak badani doświadczają substytucji, ale również na ile nabycie nowych charakterystyk fenomenalnych jest dla nich istotne. Jeden z superużytkowników substytucji sensorycznej – Neil Harbisson (2019) – opisuje, że nie tylko wykorzystuje substytucję sensoryczną w życiu codziennym, ale również doświadcza własności fenomenalnych substytucji w snach (a więc w sytuacji, gdy urządzenie substytucji sensorycznej nie wpływa na jego doświadczenie sensoryczne). Przypadek ten wskazuje, że substytucja sensoryczna może stać się immanentną charakterystyką naszego doświadczenia zmysłowego. W takiej sytuacji na pewno osoby niewidome chętniej wykorzystywałyby urządzenia substytucji w działaniu.

BIBLIOGRAFIA

- Auvray, M., & Myin, E. (2009). Perception with compensatory devices: from sensory substitution to sensorimotor extension. *Cognitive Science*, 33(6), 1036–1058. DOI: 10.1111/j.1551-6709.2009.01040.x.
- Bach-y-Rita, P., & Kercel, S.W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7(12), 541–546.
- Block, N. (2003). Tactile sensation via spatial perception. *Trends in Cognitive Sciences*, 7(7), 285–286.
- Farmer, L.W. (1978). Mobility devices. *Bulletin of Prosthetic Research*, 47–118.
- Froese, T., McGann, M., Bigge, W., Spiers, A., & Seth, A.K. (2012). The Enactive Torch: A New Tool for the Science of Perception. *IEEE Transactions on Haptics*, 5(4), 365–375.
- Harbisson, N. (2019). Hearing colors: My life experience as a cyborg. *W: Creativity, Imagination and Innovation. Perspectives and Inspirational Stories* (s. 117–125).
- Heimler, B., & Amedi, A. (2020). Are critical periods reversible in the adult brain? Insights on cortical specializations based on sensory deprivation studies. *Neuroscience & Biobehavioral Reviews*, 116, 494–507.
- Kałwak, W., Reuter, M., Łukowska, M., Majchrowicz, B., & Wierzchoń, M. (2018). Guidelines for quantitative and qualitative studies of sensory substitution experience. *Adaptive Behaviour*, 26(3), 111–127. DOI: 10.1177/1059712318771690.
- Łukowska, M., Kałwak, W., Osiński, D., Janik, J., & Wierzchoń, M. (2022). How to teach a visually impaired person to hear colours? Multidimensional training for a colour-to-sound sensory substitution device – design and evaluation. *International Journal of Human – Computer Studies*. DOI: 10.1016/j.ijhcs.2022.102925.
- Łukowska, M., Osiński, D., Bizoń-Angov, P., Ciupińska, K., Hat, K., Gwiazdziński, P., del Pin, S.H., Kałwak, W., Korczyk, M., Reuter, M., Szwed, M., & Wierzchoń, M. (submitted). Better act than see: individual differences in sensorimotor contingencies and (meta)cognitive strategies while using a colour-to-sound sensory substitution device. *SSRN*. DOI: 10.2139/ssrn.4113324.
- Meijer, P.B. (1992). An experimental system for auditory image representations. *IEEE transactions on biomedical engineering*, 39(2), 112–121.
- Osiński, D., Łukowska, M., Bizoń-Angov, P., Hjelme, D.R., & Wierzchoń, M. (2021). Colorophone 2.0: A wearable color sonification device

- generating live stereo-soundscapes – design, implementation and usability audit. *Sensors*, 21, 7351. DOI: 10.3390/s21217351.
- Proulx, M.J. (2010). Synthetic synaesthesia and sensory substitution. *Consciousness and cognition*, 19(1), 501–503.
- Poirier, C., De Volder, A.G., & Scheiber, C. (2007). What neuroimaging tells us about sensory substitution. *Neuroscience & Biobehavioral Reviews*, 31(7), 1064–1070. DOI: 10.1016/j.neubiorev.2007.05.010.
- Ptito, M., Moesgaard, S.M., Gjedde, A., & Kupers, R. (2005). Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain: A Journal of Neurology*, 128(Pt 3), 606–614. DOI: 10.1093/brain/awh380.
- Reuter, M., Kałwak, W., Sznajder, M., & Wierzchoń, M. (w recenzji). How to measure sensory substitution experience? Enactive Torch studies with qualitative methods. *Analiza i Egzystencja*.

Piotr Konderak

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

<https://orcid.org/0000-0002-9161-4870>

<https://doi.org/10.35765/slowniki.470>

Semiotyka kognitywna: zwrot ku znaczeniu

Streszczenie

DEFINICJA POJĘCIA: Semiotyka kognitywna (*cognitive semiotics*) to transdyscyplinarne badania nad strukturami znaczeń i procesami tworzenia znaczeń (*meaning-making*), łączące pojęcia i metody semiotyki, kognitywistyki i językoznawstwa, w istotny sposób odwołujące się do fenomenologii.

ANALIZA HISTORYCZNA POJĘCIA: Semiotyka kognitywna wyłoniła się jako efekt współpracy semiotyków, językoznawców, kognitywistów i filozofów poszukujących odpowiedzi na pytanie o istotę znaczenia. Punktem wyjścia były pytania o semiotyczną perspektywę na poznanie i poznawczą perspektywę na procesy semiozy.

UJĘCIE PROBLEMOWE POJĘCIA: Semiotyka kognitywna staje przed problemem sformułowania ram pojęciowych, pozwalających na jednolite ujęcie zjawisk będących przedmiotem dyskusji w semiotyce, językoznawstwie i kognitywistyce. Kluczowe pojęcie, znaczenie, rozumiane jest w terminach dynamicznej relacji pomiędzy ucieleśnionym podmiotem a środowiskiem, wyznaczonej wartością, jaką stanowią elementy środowiska dla tego podmiotu. Rozumienie takie stanowi podstawę koncepcji wielopoziomowego znaczenia, tj. tzw. hierarchii semiotycznej obejmującej cztery poziomy: życia, świadomości, użycia znaków i języka. W ujęciu metodologicznym semiotyka kognitywna opiera się na idei feno-metodologicznej triangulacji, postulującej wykorzystanie w badaniach nad strukturami znaczenia i procesami tworzenia znaczeń metod pierwszo-

drugo- i trzecioosobowych z uznaniem epistemologicznego pierwszeństwa metod pierwszoosobowych. Rozstrzygnięcie problemu relacji pomiędzy semiotyką kognitywną a kognitywistyką wymaga odwołania do enaktywnych i ucieleśnionych koncepcji poznania.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Semiotyka kognitywna dostarcza obrazu podmiotowości wykraczającego poza koncepcję przetwarzającego informacje agenta. Zaproponowane ramy pojęciowe, m.in. Hierarchia Semiotyczna, pozwalają na spójne ujęcie licznych codziennych aktywności podmiotów poznających w kategoriach procesów tworzenia znaczeń. Semiotyka kognitywna wciąż doświadcza ograniczeń związanych z integracją metod różnych podejść.

Słowa kluczowe: semiotyka kognitywna, hierarchia semiotyczna, triangulacja feno-metodologiczna, enaktywizm, ucieleśnienie

Definicja pojęcia

Semiotyka kognitywna (*cognitive semiotics*) charakteryzowana jest jako transdyscyplinarne badania nad tworzeniem znaczeń (*meaning-making*), łączące pojęcia i metody semiotyki, kognitywistyki i językoznawstwa, w istotny sposób odwołujące się do fenomenologii. Semiotykom kognitywnym przyświecała idea integracji nauk humanistycznych, społecznych i przyrodniczych w dążeniu do wyjaśnienia zjawiska znaczenia i procesów tworzenia znaczeń. Filozofia (Gottlob Frege, Hilary Putnam, Maurice Merleau-Ponty), semiotyka (Charles Sanders Peirce, Ferdinand de Saussure) i językoznawstwo (George Lakoff, Ronald Langacker, Eleanor Rosch) zapewniły podłoże teoretyczne dociekań nad procesami tworzenia znaczeń i nad strukturami znaczeniowymi. Psychologia i neuronauka – jako elementy składowe badań kognitywistycznych – pozwalały na empiryczne ujęcie interakcji podmiotu ze światem i innymi podmiotami. Procesy tworzenia znaczeń ujmowane są jako dynamiczne, zmieniające się w perspektywie ontogenetycznej i wyłaniające się w toku ewolucji.

Nazwa „semiotyka kognitywna” może sugerować, że omawiany obszar badań stanowi jedno z podejść w ramach semiotyki, rozumianej jako „teoria znaków” lub „badania nad systemami znakowymi” (Żegleń, 2000). Semiotycy kognitywni przyjmują jednak szersze rozumienie terminu „semiotyka”, charakteryzując ją jako badania nad strukturami znaczeń i procesami tworzenia znaczeń. Znaki – charakteryzowane w ramach semiotyki rozumianej klasycznie – stanowią jedynie jeden z rodzajów struktur znaczeniowych. Semioza – rozumiana w semiotyce kognitywnej jako proces tworzenia znaczeń – obejmuje szereg zjawisk wykraczających poza tworzenie znaków, posługiwanie się nimi i ich interpretację. Szersze ujęcie semiotyki i semiozy znajduje swój wyraz w postaci ramy pojęciowej zwanej hierarchią semiotyczną, określającej różne poziomy tworzenia znaczeń.

Określenie „kognitywna” z kolei może sugerować, że semiotyka kognitywna jest jednym z obszarów badań w ramach kognitywistyki, na wzór psychologii kognitywnej czy językoznawstwa kognitywnego. Relacja między oboma obszarami badań jest jednak bardziej złożona. Semiotyka kognitywna wyrosła z niezadowolenia z charakterystycznej dla klasycznej kognitywistyki (tzw. kognitywizmu) redukcji podmiotów

poznawczych do ich funkcjonalnego wymiaru. Za niewystarczające uznaje się traktowanie podmiotów poznających jako „agentów”, tj. systemów realizujących funkcje w określonym środowisku w odpowiedzi na dane wejściowe przy odpowiedniej wewnętrznej organizacji funkcjonalnej. Innymi słowy semiotycy kognitywni ujmują procesy poznawcze nie tylko w kategoriach przetwarzania informacji, ale raczej w kategoriach procesów tworzenia znaczeń i wyłaniających się w ich efekcie struktur znaczeniowych. W konsekwencji semiotyka kognitywna odwołuje się do wybranych pojęć i metod standardowej kognitywistyki, w znacznie większym stopniu opierając się jednak na tzw. podejściach ekologicznych (*4E cognitive science*) ze szczególnym uwzględnieniem podejścia enaktywnego i ucieleśnionego.

Ze względu na swoją transdyscyplinarną naturę semiotyka kognitywna nie powinna być traktowana jako dyscyplina badawcza. Problematyka tworzenia znaczeń nie tylko przekracza granice poszczególnych dyscyplin czy dziedzin badawczych, wykracza również poza aktywności akademickie, czerpiąc z doświadczeń praktyków. Tym, co łączy akademickie i pozaakademickie aktywności w ramach badań, jest wspólny cel: odpowiedź na pytanie o procesy tworzenia znaczeń.

Analiza historyczna pojęcia

Semiotyka kognitywna wyłoniła się jako efekt współpracy semiotyków, językoznawców, kognitywistów i filozofów realizujących program badawczy *Center for Cognitive Semiotics* na Uniwersytecie w Lund. W 2011 roku odbyła się 7. konferencja *Nordic Association for Semiotic Studies* pod hasłem *Towards Cognitive Semiotics. A semiotic perspective on cognition – A cognitive perspective on semiosis*. Dyskutowana w trakcie konferencji idea powiązania badań semiotycznych, językoznawczych i kognitywistycznych nie jest nowa, jej źródła można szukać w rozważaniach wybranych semiotyków. Göran Sonesson w swojej monografii *Pictorial Concepts* (1989) przyjmuje perspektywę semiotyczno-kognitywistyczną, budując siatkę pojęciową umożliwiającą analizę zjawiska ikoniczności. W szczególności sformułował on koncepcję znaku, czerpiąc z pism Husserla i badań Piageta, gdzie kluczowym elementem jest „podwójna asymetria pomiędzy dwoma częściami: jedna

z nich, wyrażenie, jest bardziej bezpośrednio doświadczana; druga część, treść, bardziej skupia naszą uwagę” (Sonesson, 2010, s. 24–25). Thomas Daddesio – uznawany za nieformalnego ojca duchowego semiotyki kognitywnej – przedstawia propozycję „kognitywnej teorii symboli” (Daddesio, 1995, s. 2) i optuje za „kognitywnym podejściem do semiozy” (Daddesio, 1995, s. 10). W polskiej tradycji badawczej Jerzy Pelc (2013) opowiada się za ściślejszym powiązaniem badań nad poznaniem i komunikacją z badaniami prowadzonymi w ramach semiotyki. Umberto Eco (1999) – jak deklaruje – nie miałby nic przeciwko ujęciu semiotyki w powiązaniu z naukami o poznaniu, nie przesądzając jednocześnie, czy semiotyka jest jednym z podejść w ramach kognitywistyki, czy też kognitywistyka jest jedną z gałęzi semiotyki.

Historycznych źródeł semiotyki kognitywnej należy szukać nie tylko w obrębie semiotyki, ale także w dwu innych obszarach badań. Wyjaśnienia procesów tworzenia znaczeń wymagają odwołania się do pojęcia podmiotu. Ideę „funkcjonalnie równoważnych” agentów – agentów realizujących tę samą funkcję niezależnie od ich fizycznej konstytucji – zastępuje pojęcie ucieleśnionego podmiotu wchodzącego w interakcje z otoczeniem (fizycznym, społecznym), które jest dla podmiotu znaczące. W tym kontekście szczególnego znaczenia nabiera publikacja Francesca Vareli, Evana Thompsona i Eleanor Rosch: *The Embodied Mind: Cognitive Science and Human Experience* – stanowiąca inspirację dla perspektyw fenomenologicznych i wprowadzająca enaktywistyczne i ucieleśnione ramy myślenia o poznaniu. Odwołanie do dynamiki procesów tworzenia znaczeń z kolei inspirowane jest podejściami ewolucyjnymi i rozwojowymi. Szczególną rolę w kształtowaniu myślenia o ewolucji tworzenia znaczeń odegrali Merlin Donald (*Origins of the Modern Mind*, 1991) oraz Terry Deacon (*The Symbolic Species: The Co-evolution of Language and the Brain*, 1997). W perspektywie rozwojowej na badania semiotyków kognitywnych znaczący wpływ miały idee Lwa Wygotsky’ego i Jeana Piageta, idee Michaela Tomasella (*Origins of Human Communication*, 2008) czy też badania Colwyna Trevarthena i Ulrike Lüdtkke nad etapami wyłaniania się intersubiektywności u dzieci.

Na tym tle konstytuowała się semiotyka kognitywna jako podejście zinstytucjonalizowane. W 2013 powstaje International Association for Cognitive Semiotics łączące badaczy zainteresowanych problematyką tworzenia znaczeń i odpowiedzialne za organizację kolejnych

konferencji: w Lund (2014), w Lublinie (UMCS, 2016), w Toronto (2018), Aachen (2022). Piąta edycja odbyła się w 2024 roku na Uniwersytecie w Lund. Od 2007 roku badacze mogą publikować w czasopiśmie *Cognitive Semiotics* (wydawany przez wydawnictwo de Gruyter). Semiotyka kognitywna wkracza również w obszar edukacyjny: bądź jako kierunek studiów, bądź jako realizowany w ramach studiów kurs.

Ujęcie problemowe pojęcia

Punktem wyjścia badań w zakresie semiotyki kognitywnej są dość ambitnie sformułowane cele. Tak jak każde inter-, multi- czy transdyscyplinarne podejście, tak i semiotyka kognitywna staje przed problemem integracji wyników badań w ramach znacząco odmiennych dyscyplin. Odmienność siatek pojęciowych oraz metodologii poszczególnych dyscyplin nie stanowi jedynego wyzwania dla badaczy. Szczególną trudność sprawia tutaj kluczowe pojęcie – znaczenie. W konsekwencji semiotyka kognitywna powinna dostarczyć dostatecznie ogólnych ram pojęciowych, które tworzą wspólną płaszczyznę dla rozstrzygnięć dotyczących znaczenia.

Wielopoziomowość znaczenia: hierarchia semiotyczna

U podstaw badań semiotyki kognitywnej legło przekonanie o wielowymiarowej naturze znaczenia, znajdującej swoje odzwierciedlenie w wieloznaczności terminu „znaczenie”. Próbując skonstruować jednolitą ramę badań nad znaczeniem i tworzeniem znaczeń, Zlatev (2009), a następnie Zlatev i Konderak (2023) zaproponowali tzw. hierarchię semiotyczną – ramę pojęciową pozwalającą ująć rozmaitego rodzaju znaczenia i różne procesy tworzenia znaczeń: znaczenie biologiczne, fenomenalne, kulturowe (choć nie językowe) i znaczenie językowe. Punktem wyjścia jest ujęcie znaczenia w kategoriach dynamicznych relacji pomiędzy podmiotem a środowiskiem, w którym ten podmiot jest osadzony, wyznaczonych wartością, jaką stanowią elementy środowiska dla tego podmiotu. Innymi słowy tworzenie znaczeń to zjawisko relacyjne, na które składają się oparte na wartościach interakcje pomiędzy podmiotem a jego

„światem”. Takie ujęcie przyniosło rozumienie znaczenia w kategoriach warstw lub poziomów znaczenia, gdzie kolejne warstwy: życia (*animation*), świadomości, użycia znaków i języka wyłaniają się, ustanawiając swoistą hierarchię poziomów tworzenia znaczeń. Zależność pomiędzy kolejnymi warstwami ma charakter dynamicznej, dwustronnej relacji. Z jednej strony, każdy kolejny poziom wymaga istnienia poziomu niższego i jest na nim ufundowany. Aktywność żywego organizmu jest niezbędna do wyłonienia się świadomości, świadomość do użycia znaków, zdolności semiotyczne stanowią podstawę tworzenia znaczeń językowych. Z drugiej strony, kolejny poziom wykracza poza („sublimuje”) poziom, na którym jest ufundowany. Modelem dla relacji pomiędzy warstwami jest Merleau-Ponty’ego (1962) relacja *Fundierung*. Ta dwukierunkowa relacja podkreśla pierwszeństwo tego, co fundujące (*originator*), w stosunku do tego, co fundowane (*originated*), w takim sensie, że to, co fundowane, jest określoną lub wyraźną formą tego, co fundujące. Jednocześnie to, co fundujące, manifestuje się w tym, co ufundowane (Merleau-Ponty, 1962, s. 458).

Najniższa warstwa w hierarchii semiotycznej, lokująca się powyżej tzw. semiotycznego dolnego progu (por. np. Nöth, 2001), wyłania się w aktywności żywych organizmów. Innymi słowy, tylko istoty ożywione zdolne są do tzw. wsobnego (*intrinsic*) znaczenia. Semioza zachodząca powyżej dolnego progu semiotycznego jest przedmiotem biosemiotyki (Jakob von Uexküll, Kalevi Kull). Ucieleśniony podmiot, jakim jest żywy organizm, działa w środowisku, którym jest Umwelt. Wartością nadającą znaczenie elementom środowiska jest przetrwanie organizmu.

Życie stanowi warunek wyłonienia się kolejnej warstwy: poziomu doświadczenia lub subiektywności. Podmioty nadające znaczenie na poziomie drugim są zdolne nie tylko do wykrywania warunków sprzyjających bądź niesprzyjających przetrwaniu oraz podejmowania stosownych działań, ale również do świadomego doświadczenia świata. Podmiot doświadczający jest intencjonalnie nakierowany na doświadczane obiekty. Na tym poziomie wyłania się znaczenie fenomenalne, zaś podmiot osadzony jest w Husserlowskim *Lebenswelt*, świecie-jako-doświadczanym. Innymi słowy to, co jest znaczące, na tym poziomie jest doświadczane.

Świadomość (fenomenalna) jest koniecznym, choć niewystarczającym warunkiem trzeciego poziomu tworzenia znaczeń, na którym powstaje

funkcja znakowa, wylaniająca się w procesach użycia i tworzenia intersubiektywnie dostępnych znaków. Znak, w myśl rozwijanych w semiotyce kognitywnej koncepcji, pojawia się, gdy podmiot dokonuje odróżnienia znaku samego, formy, jaką przyjmuje znak (np. gestu wskazującego, obrazu, symbolu) i tego, co znak reprezentuje. Podmiot na trzecim poziomie tworzenia znaczeń jest zdolny do powiązania wyrażenia (*expression*) i jego treści, opierając się na relacjach ugruntowania: wybrane zjawiska kulturowego Lifeworld są dla podmiotu znaczące dzięki relacji bliskości (wskazniki), podobieństwa (ikony) lub konwencji (symbole).

Na poziomie czwartym podmiot dąży do używania symboli w sposób poprawny i systematyczny. Oznacza to, że jest oparty na zbiorze przyjętych w danej społeczności językowej norm. Środowisko, którym osadzony jest podmiot, stanowi „uniwersum dyskursu”: codzienne konwersacje, powieści, poematy, mity, teorie naukowe i inne mówione lub pisane „teksty”. Język wymaga nie tylko zbioru symboli, ale również gramatyki – rozumianej jako system norm semiotycznych.

Podmioty zdolne do tworzenia znaczeń na wyższych poziomach są również zdolne do tworzenia znaczeń na poziomach niższych. Podmiot oceniający publikacje naukowe na temat kryzysu klimatycznego jest w stanie nie tylko rozumieć język tekstu (poziom IV hierarchii semiotycznej), ale również potrafi zinterpretować wskazania termometru (poziom III hierarchii semiotycznej) jako znak warunków atmosferycznych, zdolny jest do doświadczenia świata (wysokie temperatury bywają nieprzyjemne, poziom II hierarchii semiotycznej), ale również do regulacji swoich zachowań, by utrzymać się przy życiu (poszukiwanie schronienia przed słońcem, poziom I hierarchii semiotycznej). Na każdym z poziomów mamy do czynienia ze znaczeniami i odpowiednimi procesami tworzenia znaczeń: wysokość temperatury jest znacząca ze względu na przetrwanie organizmu; organizm, regulując swoje zachowanie, nadaje sens temu czynnikowi środowiskowemu. Sposób świadomego doświadczenia temperatury stanowi znaczenie na poziomie II, świadomość pozwala nadać sens temu doświadczeniu. Na poziomie III wskazanie termometru ma znaczenie jako znak. Interpretacja tego znaku w kategoriach innych znaków stanowi proces tworzenia znaczeń.

Hierarchia semiotyczna dostarcza ogólnych ram dla dyskusji nad procesami tworzenia znaczeń i strukturami znaczeń. W faktycznej praktyce badawczej badania nad procesami tworzenia znaczeń odwołują się

bądź do poszczególnych poziomów (na czym polega świadome doświadczanie spontanicznych gestów? Jakie gesty mają charakter ikoniczny, jakie wskaźnikowy? W jakiej relacji pozostaje język migowy do pantomimy?), bądź do relacji pomiędzy poziomami (czy – a jeżeli tak, to w jaki sposób – z aktywności pantomimicznej wyłania się język migowy?). Hierarchia semiotyczna podlega zmianom. Przykładem ilustrującym ten proces jest najnowsze wcielenie hierarchii semiotycznej (Zlatev & Konderak, 2023), gdzie poziom II rozwarstwia się na poziomy subiektywnego i intersubiektywnego tworzenia znaczeń. Ten ostatni uwzględnia zjawisko współdzielonej intencjonalności (*shared intentionality*).

Dynamika znaczeń: struktury vs. procesy

Pierwotnie, pod wpływem enaktywistycznych i ucieleśnionych ujęć podmiotów tworzących znaczenia, semiotycy kognitywni podkreślali dynamizm procesów tworzenia znaczeń. Jednak fakt powstawania relatywnie trwałych struktur znaczeniowych, zwłaszcza na poziomach użycia znaków (np. utrwalanie się gestu „OK” czy podniesionego kciuka) i języka (np. kształtowanie się określonych struktur gramatycznych), wymusił bardziej „dialektyczne” podejście do kwestii procesów i struktur. Motywacją była Merleau-Ponty’ego idea struktury świata jako *sedymentacji*, czyli zdolności utrwalania określonych struktur, i *spontaniczności*, czyli zdolności tworzenia nowych struktur. W konsekwencji znaczenie ma dualną naturę. Można je charakteryzować zarówno jako dynamiczny proces i jako stabilną strukturę. Z jednej strony usytuowane, ucieleśnione działanie daje początek strukturom znaczenia, z drugiej obecne już struktury ograniczają podejmowane działania. Zatem na każdym ze scharakteryzowanych wyżej poziomów hierarchii semiotycznej znajdujemy dwa aspekty znaczenia: stosunkowo stabilne struktury (normy) oraz rozwijające się w czasie procesy mogące prowadzić do zmiany tych norm. Badania w ramach semiotyki kognitywnej przyjmują jedną z wymienionych wyżej perspektyw (procesualną albo strukturalną) bądź zmierzają do ustalenia zasad współistnienia procesów i struktur w faktycznych interakcjach z otoczeniem. Dynamiczne współdziałanie *sedymentacji* i *spontaniczności* pojawia się na każdym z poziomów znaczenia.

Podobnie jak w przypadku poziomów znaczenia, rozróżnienie struktur znaczenia i procesów tworzenia znaczeń dostarcza pewnej ogólnej ramy dla dyskusji nad procesami semiozy. Pojęcia te można wykorzystać zarówno przy opisie kształtowania się norm językowych, jak i przy wyjaśnieniu wyłaniania się wzorców zachowań. Znajdują swoje zastosowanie zarówno w opisie ontogenetycznym, jak i w ujęciach ewolucyjnych. Faktyczne badania w ramach semiotyki kognitywnej obejmują charakterystyki procesów (i wiążą się czasami z odrzuceniem wymiaru strukturalnego) oraz charakterystyki wyłaniających się struktur.

Triangulacja feno-metodologiczna

Semiotyka kognitywna, zajmując się problematyką znaczeń, czerpie z ram pojęciowych i metodologii trzech obszarów badań: semiotyki, językoznawstwa kognitywnego i kognitywistyki. Szerokie spektrum metod stosowanych w poszczególnych obszarach badań i w poszczególnych dziedzinach oraz odmienność tych metod stanowią wyzwania dla metodologicznego wymiaru semiotyki kognitywnej.

Głównym postulatem metodologicznym przyjmowanym w semiotyce kognitywnej jest feno-metodologiczna triangulacja, określana również terminem „triangulacja fenomenologiczna”, a wcześniej terminem „triangulacja metodologiczna”. Triangulacja fenomenologiczna postuluje, aby zjawisko znaczenia opisywać integralnie, z trzech perspektyw: pierwszo-, drugo- i trzecioosobowej. W opozycji do standardowej kognitywistyki semiotyka kognitywna uznaje epistemologiczny priorytet podejścia pierwszoosobowego, ponieważ zgodnie z perspektywą fenomenologiczną wszelka wiedza ma swoje źródło w doświadczeniu. Pojęcie triangulacji ma dwa wymiary: ontologiczny i epistemologiczny. Ten pierwszy odnosi się do natury badanych zjawisk, mianowicie danych pierwszoosobowych (w przypadku badań subiektywności) czy danych trzecioosobowych (w przypadku np. neuronauki). Wymiar epistemologiczny, z drugiej strony, ujmuje zjawisko takim, jakie przedstawia się samemu zaangażowanemu badaczowi (perspektywa pierwszoosobowa), wyłania się w interakcjach z innymi (perspektywa drugoosobowa) czy bezstronnemu, niezaangażowanemu obserwatorowi (perspektywa trzecioosobowa). Przykładami metod pierwszoosobowych

mogą być analiza fenomenologiczna lub systematyczne intuicje. Metody drugoosobowe obejmują obserwację uczestniczącą, wywiady lub intersubiektywne walidacje. Przykładami metod trzecioosobowych są eksperymenty i metody statystyczne. Feno-metodologiczna triangulacja postuluje, by – na ile natura zjawiska badanego na to pozwala – integrować trzy powyższe perspektywy na znaczenie.

Postulat powyższy pomaga zrealizować drugi komponent metodologiczny semiotyki kognitywnej, którym jest tzw. pętla pojęciowo-empiryczna. Punktem wyjścia badań są często analizy pojęciowe – charakterystyczne dla semiotyki lub językoznawstwa kognitywnego. Idea pętli pojęciowo-empirycznej wyraża współzależność pomiędzy kwestiami pojęciowymi (czym jest dane zjawisko, np. czym jest spontaniczna gestykulacja?) i zagadnieniami empirycznymi (jak dane zjawisko się manifestuje, np. jak spontaniczna gestykulacja manifestuje się w zachowaniach wykładowców?), generując kolejne kroki w postępowaniu. Rozróżnienia pojęciowe zostają odzwierciedlone w badaniach empirycznych, te z kolei mogą skutkować rekonceptualizacją.

Poznanie a tworzenie znaczeń: ekologiczna (4E) kognitywistyka

Niełatwo jest określić wzajemny stosunek semiotyki kognitywnej i kognitywistyki. Przyczyny takiego stanu rzeczy tkwią zarówno w historii semiotyki kognitywnej, jak i w założeniach obu podejść. Sytuację komplikują dyskusje nad tym, czym kognitywistyka faktycznie jest. Semiotyka kognitywna rozwijała się w wyraźnej opozycji do obliczeniowego, reprezentacjonalistycznego podejścia charakteryzującego tzw. kognitywizm. Przedmiotem szczególnej krytyki stały się funkcjonalistyczne założenia kognitywistyki, którym przeciwstawiane było podejście nawiązujące do fenomenologii. Semiotyka kognitywna wyłaniała się wraz z rozwojem alternatywnych, tzw. ekologicznych (4E) podejść do poznania. Idee poznania jako ucieleśnionych interakcji z otoczeniem, idea procesów poznawczych wkraczających poza granice mózgu i ciała motywowała dyskusje nad sposobem rozumienia procesów tworzenia znaczeń i strukturami znaczenia.

W 1991 roku ukazała się książka Vareli, Thompsona i Rosch *The Embodied Mind: Cognitive Science and Human Experience*. Monografia

ta otwiera drzwi dwu ideom konstytuującym tzw. ekologiczne podejścia do poznania: enaktywizm i ucieleśnienie. Thompson w swojej monografii z 2007 roku *Mind in Life: Biology, Phenomenology and the Sciences of Mind* w systematyczny sposób prezentuje założenia „nowej kognitywistyki”. Do tych podejść nawiązuje semiotyka kognitywna, współdzieląc z nimi przede wszystkim punkt wyjścia: do realizacji procesów tworzenia znaczeń zdolne są jedynie istoty żywe. Najniższy poziom w hierarchii semiotycznej odpowiada minimalnym wymogom formułowanym w obrębie enaktywizmu. W obu podejściach procesy tworzenia znaczeń są konstytuowane przez czynniki środowiskowe i cielesne. Dynamikę semiozy można wyjaśnić tylko, odwołując się do doświadczającego, ucieleśnionego podmiotu, ustanawiającego swoją własną dziedzinę znaczących zjawisk. Tym samym enaktywizm i perspektywa ucieleśniona stanowią poznawcze fundamenty dwu pierwszych poziomów w hierarchii semiotycznej: poziomu życia i świadomości.

Procesy tworzenia znaczeń – scharakteryzowane w hierarchii semiotycznej – opierają się na dynamicznej interakcji pomiędzy podmiotem a środowiskiem. Tym samym trzy zjawiska wychodzą na pierwszy plan: dynamika procesów tworzenia znaczeń, rola podmiotowości (*agency*) w procesach tworzenia znaczeń oraz ucieleśniona natura interakcji podmiotu i środowiska. Istoty żywe, jak to ujmuje Thompson (2007), aktywnie się generują i podtrzymują swoje istnienie. Tym samym ich działania ustanawiają to, co dla nich znaczące w toku eksploracji środowiska. Podmioty tworzące znaczenia nie są biernymi agentami oczekującymi na informacje, by je później przetworzyć. Raczej – w duchu psychologii Jamesa Gibsona – aktywnie poszukują informacji, eksplorując swoje środowisko. W konsekwencji pojmowanie podmiotu tworzącego znaczenia odbiega od rozumienia agenta w duchu kognitywizmu. Rola podmiotu nie ogranicza się do operacji na amodalnych symbolach realizowanych zgodnie z zestawem reguł. Podmiot traktowany jest jak system autonomiczny, działający celowo, mający poczucie sprawczości.

Różnicę w ujęciu podmiotowości na różnych poziomach odzwierciedla rozróżnienie pomiędzy *quasi*-podmiotami i podmiotami właściwymi. W obu przypadkach tworzą one znaczenia, jednak są to odmienne sposoby tworzenia znaczeń. *Quasi*-podmioty to istoty żywe dysponujące *quasi*-doświadczeniem, wykazujące się biologicznie zdeterminowanymi, nakierowanymi na cel działaniami. Zachowania

quasi-podmiotów mogą być wyjaśniane na pierwszym poziomie hierarchii semiotycznej.

Tylko podmioty właściwe są zdolne do subiektywnego doświadczania świata – ich aktywność wyjaśniana jest na drugim poziomie hierarchii semiotycznej. Wymiar funkcjonalny, choć niezbędny, nie jest wystarczający do zrozumienia podmiotu na tym poziomie. Doświadczenia podmiotu ustanawiają (*enact*) znaczące elementy środowiska. Doświadczenie to – zgodnie z perspektywą Merleau-Ponty’ego – ma charakter ucieleśniony: ciało jest warunkiem doświadczania. Ciało podmiotu pozostaje w dwukierunkowej relacji ze środowiskiem: z jednej strony, środowisko stale wpływa na nasze ciała i je reguluje, z drugiej strony, nasze ciała stale wpływają na otoczenie. W efekcie wyłania się wzorzec dynamicznej, ucieleśnionej interakcji.

Rola dwu pozostałych komponentów: rozszerzonego i usytuowanego poznania, jest znacznie mniej klarowna w wyjaśnieniach procesów tworzenia znaczeń. Z jednej strony, podejście rozszerzone daje się uzgodnić z enaktywizmem czy podejściem ucieleśnionym. Twierdzenie, zgodnie z którym nie można procesów tworzenia znaczeń zamykać w głowie czy nawet w ciele, wydaje się wychodzić naprzeciw idei relacyjnej natury procesów tworzenia znaczeń. Z drugiej jednak strony, funkcjonalistyczne fundamenty koncepcji Clarka i Chalmersa (1998) każą semiotykom kognitywnym poszukiwać alternatywnej koncepcji poznania rozszerzonego. Drugim problemem, pojawiającym się w kontekście poznania usytuowanego i rozszerzonego, jest pojęcie podmiotu działającego. Rozszerzone systemy tworzące znaczenia, powstające w wyniku parowania (*coupling*) podmiotu i elementów środowiska, mają zbyt „ulotną” naturę, by stanowić podstawę trwałej podmiotowości. Łatwość parowania i deparowania utrudnia – lub nawet uniemożliwia – ujęcie systemu tworzącego znaczenia jako podmiotu w sensie semiotyki kognitywnej. Ostatecznie ujęcie rozszerzone – w dyskutowanej tu wersji – nie odzwierciedla w pełni doniosłości idei ucieleśnienia. Pojęcie ciała jako narzędzia nie oddaje roli ciała procesach tworzenia znaczeń, w szczególności w ustanawianiu tego, co znaczące, w doświadczaniu świata, w przejściu od subiektywności do intersubiektywności.

Refleksja systematyczna z wnioskami i rekomendacjami

Semiotyka kognitywna docenia wielość i różnorodność naszych interakcji ze światem, dzięki którym nasze działania są nie tylko celowe czy skuteczne, ale również znaczące. Migracja ze względu na warunki klimatyczne, zmiana środowiska w poszukiwaniu przyjaciół, podążanie wzrokiem za znajomą postacią, komunikacja z wykorzystaniem gestów i komunikacja łamanym językiem obcym, spontaniczny taniec na koncercie i taniec na scenie jako element baletu, przygotowywanie smacznych potraw i przygotowywanie eleganckich dań – to tylko nieliczne z codziennych aktywności poddawanych analizom w kategoriach czynności poznawczych. Wiele z nich (a może każdą) da się opisać jako zachowania celowe, oparte na docierających do podmiotów informacjach, podlegające mniej lub bardziej określonym regułom. Jednocześnie każda z nich stanowi aktywność ucieleśnionego, odczuwającego, działającego podmiotu wchodzącego w interakcje z otoczeniem fizycznym lub społecznym. Dążąc do możliwie pełnego obrazu istot żywych, semiotycy kognitywni podejmują próby rozszerzenia idei przetwarzającego informacje agenta, uznając subiektywne doświadczenie i ucieleśnioną interakcję za nieeliminowalne elementy wyjaśnienia ludzkich działań w świecie.

Szereg badaczy zakwestionowałby powyższe stwierdzenie, domagając się raczej zastąpienia agenta standardowej kognitywistyki podmiotem tworzącym znaczenia. W tym duchu toczy się wiele dyskusji w ramach podejść ekologicznych, domagających się eliminacji podejścia reprezentacyjnego i teorio-informacyjnego i zastąpienia go podejściami enaktywnymi, ucieleśnionymi i rozszerzonymi. Pomimo początkowego entuzjastycznego (i poniekąd bezkrytycznego) przyjęcia koncepcji enaktywnych czy ucieleśnionych wśród badaczy semiotyki kognitywnej rośnie świadomość istotnych ograniczeń podejść ekologicznych (4E) i świadomość doniosłości osiągnięć standardowej kognitywistyki. Semiotyka kognitywna szuka zatem drogi umożliwiającej wieloaspektowe ujęcie podmiotów – jako działających i odczuwających, szuka „nowej nauki o umyśle”. Obiecującym podejściem przekraczającym ograniczenia funkcjonalnie rozumianej kognitywistyki a jednocześnie nieeliminującym dorobku kognitywistyki, jest koncepcja „umysłu jako

amalgamatu” (*Amalgamated mind*) rozwijana przez Rowlandsa (2010; por. też Konderak, 2018).

Przyjęte przez badaczy ramy teoretyczne są o tyle wartościowe, o ile pozwalają na wzbogacenie perspektywy w odniesieniu do szerokiego spektrum zjawisk. Na ile hierarchia semiotyczna czy triangulacja fenomenologiczna są w stanie spełnić oczekiwania badaczy? Z jednej strony, coraz liczniejsze aplikacje hierarchii semiotycznej, m.in. w analizach takich zjawisk jak: ewolucja języka, rozwój komunikacji u dzieci, polisemiotyczność komunikatów, narratywność pokazują potencjalną nośność tej ramy. Z drugiej strony, wciąż ambiwalentny stosunek części badaczy do programu standardowej kognitywistyki, sceptycyzm w odniesieniu do psychologicznej idei wewnętrznej reprezentacji, kwestionowanie neurocentryzmu w kognitywistyce pokazują trudności w zastosowaniach postulatów triangulacji. Świadectwem tych trudności są ograniczenia programu neurofenomenologii.

BIBLIOGRAFIA

- Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, 58, 10–23.
- Daddesio, T. (1995). *On Minds and Symbols: The Relevance of Cognitive Science for Semiotics*. Berlin: De Gruyter Mouton.
- Eco, U. (1999). *Semiotics in the Next Millennium*. Pozyskano z: <http://www.umbertoeco.it/CV/Semiotics%20in%20the%20next%20millennium.pdf> (dostęp: 13.07.2024).
- Konderak, P. (2018). *Mind, Cognition, Semiosis: Ways to Cognitive Semiotics*. Lublin: UMCS Press.
- Merleau-Ponty, M. (1962). *Phenomenology of Perception*. London: Routledge & Kegan Paul.
- Merleau-Ponty, M. (2001). *Fenomenologia percepcji*. Przeł. M. Kowalska i J. Migasiński, postłowie J. Migasiński. Warszawa: Fundacja Aletheia.
- Nowakowski, P., & Komendziński, T. (2014). Cognition as shaking hands with the world. Introduction. *Avant*, V(2), 10–16.
- Nöth, W. (2001). Protosemiotics and physicosemiotics. *Sign System Studies*, 29(1), 13–25.
- Pelc, J. (2013). Dwa pojęcia komunikacji. *Studia Semiotyczne*, XXVIII–XXIX, 107–119.

- Rowlands, M. (2010). *The New Science of the Mind*. Cambridge, MA: The MIT Press.
- Sonesson, G. (1989). *Pictorial Concepts: Inquiries into the Semiotic Heritage and Its Relevance to the Interpretation of the Visual World*. Lund: Lund University Press.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology and the Sciences of Mind*. London: Belknap Press.
- Varela, F.J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: The MIT Press.
- Zlatev, J. (2009). The semiotic hierarchy: life, consciousness, signs and language. *Cognitive Semiotics*, 4, 169–200.
- Zlatev, J., & Konderak, P. (2023). Consciousness and Semiosis. W: J. Pelkey (Red.), *Bloomsbury Semiotics Volume 1: History and Semiosis*. London: Bloomsbury Academic.
- Żegleń, U. (2000). *Wprowadzenie do semiotyki teoretycznej i semiotyki kultury*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.

Łukasz Wiraszka
Uniwersytet Jagielloński w Krakowie
<https://orcid.org/0000-0002-3653-5926>
<https://doi.org/10.35765/slowniki.448>

Językoznawstwo kognitywne

Streszczenie

DEFINICJA POJĘCIA: Językoznawstwo kognitywne jest jednym z nurtów współczesnej lingwistyki, obejmującym szereg komplementarnych teorii i metodologii, opartych na wspólnych założeniach dotyczących natury języka. Najbardziej fundamentalne z tych założeń mówi, że język stanowi integralną część ludzkiego poznania.

ANALIZA HISTORYCZNA POJĘCIA: Początki lingwistyki kognitywnej w szerokim rozumieniu sięgają połowy XX w. i mają związek z tzw. rewolucją kognitywną. Rozwinięta przez Noama Chomsky'ego generatywna teoria języka odnosiła się do zdolności językowej tkwiącej w umyśle człowieka, lecz w odróżnieniu od współczesnego językoznawstwa kognitywnego ujmowała ją jako mechanizm generujący gramatyczne zdania danego języka. Pionierami współczesnej lingwistyki kognitywnej byli niegdysiejsi uczniowie Chomsky'ego, którzy dostrzegając ograniczenia generatywnego podejścia do języka, stworzyli na przełomie lat 70. i 80. XX w. teoretyczne podwaliny językoznawstwa kognitywnego „drugiej generacji”.

UJĘCIE PROBLEMOWE POJĘCIA: Do kluczowych postulatów językoznawstwa kognitywnego należą: uzgadnianie teorii języka z aktualną wiedzą na temat funkcjonowania ludzkiego umysłu (tzw. zobowiązanie do kognitywizmu), prymat znaczenia w dociekaniach językoznawczych, rozszerzenie przedmiotu badań na wszelkie rodzaje użycia języka, w tym język metaforyczny, oraz zwrócenie uwagi na rolę ucieleśnienia. Lingwistyka kognitywna ujmuje metaforę jako zjawisko poznawcze (sposób myślenia) i przyjmuje subiektywistyczną koncepcję znaczenia językowego.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI:

Badając zjawiska językowe, językoznawstwo kognitywne formułuje tezy na temat procesów mentalnych. Pełna realizacja tego zadania wymaga badań interdyscyplinarnych prowadzonych w ramach nauk kognitywnych.

Słowa kluczowe: językoznawstwo kognitywne, kognitywistyka, gramatyka kognitywna

Definicja pojęcia

Językoznawstwo kognitywne jest jednym z nurtów we współczesnej lingwistyce, obejmującym szereg odrębnych, choć komplementarnych, teorii i metodologii, opartych na wspólnych fundamentalnych założeniach dotyczących natury języka oraz sposobów uprawiania językoznawstwa. Początki współczesnej lingwistyki kognitywnej sięgają przełomu lat 70. i 80. XX wieku, kiedy to – w reakcji na niektóre z tez dominującej wówczas w amerykańskim językoznawstwie generatywnej szkoły Chomsky’ego oraz pod wpływem rozwoju badań z zakresu psychologii poznawczej – sformułowano pierwsze kluczowe postulaty, które dały podwaliny pod rozwinięte w latach 80. kluczowe koncepcje teoretyczne tzw. ojców założycieli językoznawczego kognitywizmu, wśród których czołowe miejsce zajmują G. Lakoff, R. Langacker i L. Talmy.

Wraz z innymi dziedzinami, takimi jak psychologia poznawcza, filozofia, antropologia, neuronauka oraz badania nad sztuczną inteligencją, lingwistyka kognitywna wnosi swój wkład do kognitywistyki, interdyscyplinarnego programu badawczego, którego głównym przedmiotem są mechanizmy ludzkiego poznawania i myślenia. Udział językoznawstwa kognitywnego w „przedsięwzięciu” kognitywistycznym wiąże się z ujmowaniem języka jako pewnego rodzaju wiedzy, która tkwi w umysłach jego użytkowników. I choć podstawowe cele lingwistyki kognitywnej nie różnią się zasadniczo od zadań, jakie stawiają sobie językoznawcy spoza kręgu kognitywistycznego – obejmują one właściwy opis zjawisk językowych oraz sformułowanie teorii umożliwiających ich wyjaśnienie – odpowiedzi na pytania o naturę wiedzy językowej językoznawcy-kognitywiści „szukają w tym, co uznają za źródło, a zarazem rację istnienia języka: w zdolnościach ludzkiego umysłu” (Tabakowska, 1995, s. 5).

Takie podejście wynika wprost z jednego z fundamentalnych założeń językoznawstwa kognitywnego, wedle którego „język stanowi integralną część ludzkiego poznania” (Langacker, 1987, s. 12; przeł. Ł.W.). Istotnym elementem tego założenia jest podkreślenie „zintegrowania” zdolności językowej z innymi aspektami funkcjonowania ludzkiego umysłu. Składająca się na zdolność językową wiedza nie stanowi w umyśle człowieka autonomicznego komponentu pojęciowego (modułu), lecz opiera się na tych samych ogólnych procesach mentalnych co inne rodzaje

wiedzy i wchodzi w ścisłe interakcje z innymi zdolnościami poznawczymi, takimi jak kategoryzacja, pamięć, percepcja, rozumowanie czy skupienie uwagi. W językoznawstwie kognitywnym przyjmuje się, że język nie tylko wykorzystuje, ale również odzwierciedla mechanizmy ludzkiego poznawania, dlatego badanie zjawisk językowych sprowadza się w ostatecznym rozrachunku do badania mechanizmów mentalnych składających się na aparat poznawczy człowieka (Tabakowska, 1995, s. 12).

Konceptualne i metodologiczne ramy współczesnej lingwistyki kognitywnej wyznaczają jej dwa fundamentalne zobowiązania, wyłożone przez Lakoffa (1990) w pierwszym numerze czasopisma *Cognitive Linguistics*: zobowiązanie do kognitywizmu oraz obowiązek do generalizacji. Pierwsze z nich wymaga, aby wszelkie tezy odnoszące się do zjawisk językowych pozostawały w zgodzie z dotychczasowymi ustaleniami pozostałych nauk kognitywnych w zakresie funkcjonowania ludzkiego umysłu. Na przykład struktura kategorii pojęciowych przypisywanych wyrażeniom językowym powinna uwzględniać efekty prototypowe wykazane w eksperymentach psychologicznych, zaś postulowanie kategorii językowych wyłącznie na podstawie klasycznego, arystotelesowskiego modelu kategoryzacji, tzn. definiowanych przez zbiór cech koniecznych i wystarczających, wymagałoby specjalnego, empirycznego uzasadnienia (Lakoff, 1990, s. 41). Rozumiane w ten sposób zobowiązanie do kognitywizmu ma zatem maksymalizować wiarygodność psychologiczną teorii językoznawczych.

Drugie z fundamentalnych zobowiązań językoznawstwa kognitywnego, zobowiązanie do generalizacji, obliuguje do poszukiwania uogólnień wysokiego poziomu. Nie chodzi jednak o typowe dla ujęć formalnych dążenie do maksymalnej oszczędności opisu, lecz o rozpoznawanie wspólnych mechanizmów i zasad dla różnych poziomów struktury języka i odmiennych aspektów jego użycia. Skoro język jest zdolnością bazującą na ogólnych właściwościach ludzkiego umysłu, to samo dotyczy wszystkich jego elementów. Oznacza to, przykładowo, że zjawiska składniowe w dużej mierze można – i należy – wyjaśniać za pomocą tych samych mechanizmów i konstruktów pojęciowych co zjawiska semantyczne, a podział na poszczególne poziomy opisu struktury języka motywowany jest jedynie czynnikami praktycznymi (Tabakowska, 1995, s. 12).

Uznawanie języka za integralną część ludzkiego poznania prowadzi w naturalny sposób do podkreślania jego funkcji poznawczej. W takim ujęciu język jest czymś więcej niż tylko narzędziem komunikacji międzyludzkiej: będąc w istocie ustrukturyzowanym zbiorem kategorii pojęciowych, za pomocą których postrzegamy rzeczywistość i rozumiemy nasze doświadczenia, język stanowi zdolność poznawczą uczestniczącą w przetwarzaniu i organizacji wiedzy o świecie (Geeraerts & Cuyckens, 2007, s. 5). Pojmowanie języka jako swojego rodzaju magazynu informacji o świecie nakazuje z kolei poszukiwanie motywacji semantycznych we wszelkich dociekaniach językoznawczych.

Do najważniejszych koncepcji teoretycznych językoznawstwa kognitywnego należą: stworzona przez R. Langackera gramatyka kognitywna, teoria metafory pojęciowej zaproponowana przez G. Lakoffa i M. Johnsona, teoria przestrzeni mentalnych i integracji pojęciowej sformułowana przez G. Fauconniera i M. Turnera oraz różne odmiany gramatyki konstrukcji, zaprezentowane m.in. przez Ch. Fillmore'a, W. Crofta i A. Goldberg.

Analiza historyczna pojęcia

Kognitywistyczny nurt w językoznawstwie został *de facto* zapoczątkowany przez amerykańskiego lingwistę N. Chomsky'ego, który – wraz z kilkoma pionierami badań nad sztuczną inteligencją – uczestniczył w połowie lat 50. XX wieku w tzw. (pierwszej) rewolucji kognitywnej (określanej też mianem zwrotu kognitywnego), polegającej na odrzuceniu behawioryzmu (zob. np. Chomsky, 1959) i skierowaniu wysiłków badawczych na funkcjonowanie ludzkiego umysłu z zamiarem stworzenia takiego opisu zachodzących w nim procesów, aby można je było symulować za pomocą komputerów (por. Bremer, 2023, s. 74–75). Rozpoczęta przez Chomsky'ego rewolucja w językoznawstwie obejmowała odejście od typowego dla amerykańskiej wersji strukturalizmu (deskrytywizmu) czysto opisowego podejścia do języka jako abstrakcyjnego, statycznego systemu, tzn. zbioru struktur językowych i zachodzących między nimi relacji, na rzecz formułowania tez składających się w istocie na teorię ludzkiego umysłu, a konkretnie tej części umysłu, która odpowiada za język. Z tego powodu zainicjowany przez Chomsky'ego

nurt generatywny zaliczyć należy do językoznawstwa kognitywnego w szerokim rozumieniu (Taylor, 2007, s. 5–7).

Jednym z założeń leżących u podstaw generatywizmu był autonomiczny charakter zdolności językowej człowieka. Zgodnie z tym podejściem język stanowi w umyśle człowieka odrębny, rządzący się swoimi prawami moduł, a wiedza językowa jest odmiennym od innych, szczególnym rodzajem wiedzy nabytej w wyniku działania unikalnego, w dużej mierze wrodzonego mechanizmu przyswajania języka (Chomsky, 1982). Przyjęcie założenia o autonomii języka zapewniło Chomsky’emu niezależność językoznawstwa od innych dziedzin nauki, których przedmiotem zainteresowania jest ludzki umysł, a tym samym pozwoliło na postulowanie dowolnych w gruncie rzeczy mechanizmów na wyjaśnienie zjawisk językowych.

Wpisując się w popularną w tamtym okresie tendencję do postrzegania ludzkiego umysłu jako binarnej maszyny liczącej (Bremer, 2023, s. 75), Chomsky stworzył koncepcję języka o charakterze „operacyjnym”, na którą składał się przypominający algorytm zbiór reguł gramatycznych opisujących działania na strukturach językowych (Heinz, 1983, s. 405–406). Ujmowana w ten sposób gramatyka miała umożliwiać „wygenerowanie” (stąd nazwa tej teorii) wszystkich poprawnych i tylko poprawnych (tzn. gramatycznych) zdań danego języka. Aby zapewnić algorytmiczną precyzję, reguły gramatyczne musiały być ściśle sformalizowane, najlepiej na wzór reguł arytmetyki, w związku z czym poszczególne elementy gramatyki generatywnej przybierały postać rekursywnie stosowanych operacji na (nierzadko pustych semantycznie) symbolach. Teorię Chomsky’ego cechował więc daleko posunięty formalizm, jak również przejęte od nauk ścisłych dążenie do maksymalnej oszczędności opisu, tak aby za pomocą jak najmniejszej liczby reguł można było wygenerować jak największą liczbę struktur składniowych.

Konsekwencją ujmowania zdolności językowej w kategoriach sformalizowanego mechanizmu generującego zdania gramatyczne był prymat formy nad znaczeniem (pierwsze wersje gramatyki generatywnej były wręcz całkowicie antysemaniczne) oraz uznanie poprawności generowanych struktur składniowych za kwestię fundamentalną (Heinz, 1983, s. 402, 407). Z tego powodu najważniejszą rolę w generatywnej teorii Chomsky’ego odgrywała składnia – wyróżniana zresztą jako jeden z autonomicznych „podmodułów” języka – zaś wymóg poprawności

skłonił amerykańskiego lingwistę do przeprowadzenia wyraźnej linii demarkacyjnej pomiędzy kompetencją językową (*competence*) a tym, jak ludzie w rzeczywistości posługują się językiem, czyli wykonaniem (*performance*). Stanowiąca według Chomsky'ego właściwy przedmiot rozważań językoznawczych i, wobec tego, modelowana przez gramatykę generatywną kompetencja oznaczała zdolność językową wyimaginowanego idealnego użytkownika języka, który posługuje się wyłącznie zdaniami gramatycznymi, podczas gdy codzienne wypowiedzi rzeczywistych użytkowników języka – obfitujące we wszelkiego rodzaju niegramatyczności oraz zjawiska niedające się podporządkować regułom – relegowano do domeny wykonania i jako takie wyrzucono z zakresu zainteresowań językoznawców. Skupienie na kryterium poprawności oraz hołdowanie zasadzie konstruowania teorii z matematyczną elegancją doprowadziły ostatecznie do tego, że kolejne wersje gramatyki generatywnej były w stanie opisać jedynie niewielki procent struktur pojawiających się w rzeczywistej komunikacji językowej (Tabakowska, 1995, s. 10).

Oderwanie generatywistów od autentycznych danych językowych stanowiło jeden z przejawów – motywowanego kartezjańskim dualizmem – usuwania poza krąg zainteresowań językoznawczych wszelkich czynników wynikających z fizycznej natury człowieka. Jak to ujmuje E. Tabakowska, dla generatywistów

przedmiotem zainteresowania językoznawcy jest kompetencja językowa idealnego użytkownika języka – wolnego od wszelkich fizycznych i psychicznych niedoskonałości, zawsze posłusznego regułom gramatyki, kierującego się rozsądkiem zamiast emocjami, konsekwentnego i w swoich językowych poczynaniach odpornego na wpływy pozajęzykowego kontekstu wypowiedzi – jednym słowem pozbawionego tego złożonego „czynnika ludzkiego”, który decyduje o indywidualności każdego z nas (Tabakowska, 1995, s. 8).

Owego „czynnika ludzkiego” pozbawiony był w teorii Chomsky'ego nawet ten jej komponent, który odpowiadał za semantyczną warstwę języka. O ile gramatyka generatywna z okresu pionierskiej rozprawy Chomsky'ego z 1957 r. pt. *Syntactic structures* w ogóle nie obejmowała elementów odpowiadających za analizę znaczenia generowanych konstrukcji składniowych (co ilustruje ukute przez Chomsky'ego słynne zdanie *Colorless green ideas sleep furiously* 'Bezbarwne zielone idee

śpią wściekle', którego kluczową właściwością miała być jego gramatyczność), o tyle w pierwszej połowie lat 60., wraz z rozwojem następnej jej wersji, tzw. gramatyki transformacyjno-generatywnej, nastąpił w generatywizmie swoisty zwrot semantyczny (Heinz, 1983, s. 402). Choć jednak w kolejnych koncepcjach komponent semantyczny odgrywał coraz bardziej istotną rolę, była to semantyka oparta na logice formalnej i warunkach prawdziwości, a więc odnosząca się do (modelu) świata obiektywnego, oderwanego od człowieka będącego użytkownikiem języka i właściwego mu sposobu postrzegania rzeczywistości (tzw. obiektywistyczna teoria znaczenia – zob. Johnson, 1987, s. xxii i nast., a także Lakoff, 2011, rozdz. 11). W dodatku transformacyjno-generatywny model języka przewidywał czysto formalne operacje na wygenerowanych bazowych strukturach składniowych (tzw. transformacje), których zastosowanie prowadziło do konstruowania zdań o odmiennej strukturze składniowej, posiadających jednak tę samą logiczną strukturę semantyczną co zdanie wyjściowe. Oznaczało to, że wybór określonej struktury składniowej przez użytkownika języka nie wpływa na znaczenie wypowiedzi językowych, o ile odnoszą się one do tego samego stanu rzeczy w rzeczywistości obiektywnej.

Największe trudności sprawiała Chomsky'emu i jego zwolennikom metafora. Okazało się bowiem, że bez względu na stopień skomplikowania i formalizacji postulowane w ramach ich teorii operacje nie były w stanie wygenerować wszechobecnych w wypowiedziach językowych wyrażen metaforycznych. Zadanie to utrudniał dodatkowo przyjęty dla modułu semantycznego aparat pojęciowy z zakresu logiki formalnej, bazujący na wartościach prawdy i fałszu. Przenośne użycie języka stanowiło więc dla generatywistów „ewidentny przypadek pogwałcenia reguł gramatycznych i w ramach tworzonego modelu można [go] było potraktować jedynie jako zjawisko nietypowe i marginalne: chimeryczny ornament zaczepiony u obrzeży «normalnej» tkanki języka” (Tabakowska, 1995, s. 9). Z tego powodu metaforę zwykle pomijano, uznając ją za zjawisko nieinteresujące z punktu widzenia zadań teorii językoznawczej. Z drugiej jednak strony, teoria generatywna osiągnęła taki stopień abstrakcji, że przewidywała operacje i obiekty, których nie można było zidentyfikować w żadnym materiale językowym (Taylor, 2007, s. 8). Postulowanie „niewidzialnych” elementów, takich jak ślady czy kategorie puste, było oczywiście nastawione na zwiększenie zakresu możliwych

do wygenerowania poprawnych struktur składniowych danego języka. W ten sposób gramatyka transformacyjno-generatywna niejako narzucała ogólnej teorii umysłu unikalne dla języka procesy i obiekty mentalne niezajdujące bezpośredniego potwierdzenia ani w konkretnych danych językowych, ani w ustaleniach pozostałych nauk kognitywnych.

Generatywną szkołę Chomsky'ego, jako wczesną postać językoznawstwa kognitywnego w szerokim rozumieniu, cechowały zatem: modularność, formalizm, prymat składni oraz swojego rodzaju racjonalizm, przejawiający się skupieniem wysiłków badawczych na wyidealizowanej kompetencji w oderwaniu od posługującego się nią człowieka i wytwarzanych przez niego danych językowych. Tendencje te i wynikające z nich ograniczenia skłoniły niektórych uczniów i współpracowników Chomsky'ego, takich jak R. Langacker czy G. Lakoff, do odrzucenia podejścia generatywnego i podjęcia prób wypracowania własnych koncepcji językoznawczych. Stworzona w rezultacie przez Langackera gramatyka kognitywna oraz zaproponowana przez Lakoffa i współpracującego z nim filozofa M. Johnsona teoria metafory pojęciowej należą do najważniejszych koncepcji współczesnej lingwistyki kognitywnej, stanowiącej część kognitywistyki „drugiej generacji” (Bremer, 2023, s. 76).

Ujęcie problemowe pojęcia

Postulaty leżące u podstaw koncepcji teoretycznych Lakoffa, Langackera i innych pionierów językoznawstwa kognitywnego takiego, jakim jest ono dzisiaj, obejmowały przede wszystkim: odrzucenie formalizmu i dogmatu o autonomii języka, przeniesienie punktu ciężkości ze składni na semantykę, włączenie w spektrum analizowanych zjawisk języka figuratywnego oraz zwrócenie uwagi na rolę ucieleśnienia w rozwoju aparatu poznawczego, a w konsekwencji również zdolności językowej człowieka.

Odejście od zasady modularności i wyprowadzanej z niej niezależności języka od innych zdolności poznawczych skutkuje przyjęciem wspomnianego już na początku obowiązku uzgadniania teorii lingwistycznej z dotychczasowymi ustaleniami na temat funkcjonowania ludzkiego umysłu, a także poszukiwania w strukturze języka odbicia ogólnych procesów mentalnych oraz ścisłych powiązań z innymi zdolnościami

poznawczymi. Jedną z kluczowych dla języka zdolności poznawczych jest proces kategoryzacji, czyli tworzenia kategorii pojęciowych i kwalifikowania do nich określonych elementów naszego doświadczenia. Związek zdolności językowej ze zdolnością kategoryzacji jest dość oczywisty, bowiem wyrazy danego języka stanowią punkty dostępu do struktur pojęciowych związanych z różnymi kategoriami. Co więcej, wyróżniane przez lingwistów jednostki językowe są również podlegającymi tym samym procesom kategoriami pojęciowymi. Na przykład fizycznie odmienne sygnały dźwiękowe kategoryzowane są przez użytkowników danego języka jako przykłady tego samego fonemu (por. Taylor, 2007, s. 11). Zgodnie ze zobowiązaniem do kognitywizmu, w swych językoznawczych dociekaniach dotyczących różnych aspektów języka lingwiści kognitywni wykorzystują wypracowany na gruncie psychologii poznawczej model kategoryzacji oparty na prototypie (np. Taylor, 2001).

Innym przykładem łamania barier między językiem a pozostałymi zdolnościami poznawczymi i rozpoznawania efektów ich działania w strukturze języka jest odwoływanie się do opisanego przez psychologię postaci dzielenia obiektów podlegających percepcji człowieka na figurę i tło. W wyniku tego zjawiska niektóre elementy pola percepcji przyciągają znacznie większą uwagę postrzegającego je podmiotu niż inne i wobec tego stanowią percepcyjną figurę, podczas gdy pozostałe elementy stanowią percepcyjne tło. Okazuje się, że ten mechanizm poznawczy uwiadcza się również w różnych aspektach funkcjonowania języka, a fakt ten wkomponowali w swoje teorie m.in. R. Langacker i L. Talmy. W teoriach tych zjawisko figury i tła ma charakter konceptualny: nie odnosi się do percepcji zmysłowej, lecz do dzielenia treści pojęciowych symbolizowanych w wypowiedzi językowej na wyrazistą poznawczo figurę i niewyraziste tło. I tak, w duchu wspomnianego również na wstępie zobowiązania do generalizacji, opozycja figura/tło znajduje zastosowanie zarówno w analizie zjawisk semantycznych, jak i gramatycznych. Na przykład struktura semantyczna rzeczownika *ręka* obejmuje odniesienie do całego ludzkiego ciała, stanowiącego konceptualne tło dla tej jego części, którą nazywamy za pomocą wspomnianego rzeczownika, a która wobec tego stanowi konceptualną figurę. Z kolei spośród dwóch rzeczowników w wyrażeniu *chmury nad miastem* bardziej wyrazisty pojęciowo jest wyraz *chmury* – to na nim, a precyzyjniej: na symbolizowanym przezeń obiekcie skupia swą uwagę użytkownik języka, wobec czego

obiekt ten pełni funkcję figury w odróżnieniu od *miasta*, które pełni w tym przykładzie funkcję tła. Z odwrotnym układem figura/tło będziemy mieć jednak do czynienia w przypadku frazy *miasto pod chmurami*, gdzie bardziej wyrazistym elementem (figurą) jest rzeczownik *miasto*. Zauważmy przy tym, że obiektywnie rzecz biorąc, obydwa wyrażenia odnoszą się do tej samej sytuacji. W podobny sposób gramatyczny podmiot zdania stanowi figurę w porównaniu ze stanowiącym tło dopełnieniem, w związku z czym (odnoszące się do obiektywnie tej samej sytuacji) zdania *Policja zatrzymała złodzieja samochodów* oraz *Złodziej samochodów został zatrzymany przez policję* odzwierciedlają odmienne układy figura/tło, przez co odmienna jest też ich interpretacja semantyczna.

Towarzyszące rozwojowi współczesnej lingwistyki kognitywnej odrzucenie ideału matematyczno-logicznego formalizmu wynikało, z jednej strony, z narastającego w kognitywistyce w ogóle sceptycyzmu co do słuszności pojmowania ludzkiego umysłu na wzór operującej na symbolach maszyny Turinga, z drugiej natomiast z przekonania, że typowe dla nauk ścisłych formuły logiczne nie przystają do materii, jaką jest język naturalny, który „wymyka się ścisłym regułom logiki i na ogół bywa w ich traktowaniu niekonsekwentny” (Tabakowska, 1995, s. 11). Ponadto przyświecający kognitywistom imperatyw objęcia zakresem dociekań językoznawczych wszelkich – nawet wymykających się sztywnym regułom – zjawisk językowych uzasadnia rezygnację z dążenia do maksymalnej oszczędności opisu i ogranicza znaczenie wymogu elegancji teorii. Jako przykład posłużyć tu może fakt, że gramatyka kognitywna dopuszcza, a nawet przewiduje przypadki, w których użytkownik języka zna ogólną zasadę gramatyczną (np. regułę tworzenia form liczby mnogiej rzeczowników), a jednocześnie przechowuje w pamięci określoną formę językową będącą wynikiem zastosowania tej reguły (np. w przypadku często używanej formy liczby mnogiej, takiej jak *oczy* czy *włosy*). Sytuacja, w której gramatyka danego języka obejmuje zarówno opis zasady ogólnej, jak i konkretne rezultaty jej zastosowania, jest – z powodu oczywistej redundancji – niepożądana z punktu widzenia elegancji teorii, a typowe dla ujęć formalnych wykluczanie takich przypadków nazywa Langacker „fałszywą alternatywą reguły/listy” (Langacker, 1987, s. 29; Taylor, 2007, s. 17). Przykład ten pokazuje, że w językoznawstwie kognitywnym kryterium oszczędności opisu ustępuje wiarygodności psychologicznej.

Sprzeciw wobec formalizmu został ponadto niejako programowo wpisany przez Langackera w rozwijaną przez niego teorię języka w postaci tzw. warunku zawartości (Langacker, 1987, s. 53; 2009, s. 44). Warunek ten nakłada istotne ograniczenia co do typu jednostek językowych dopuszczalnych w kognitywnym modelu języka: mogą to być jedynie odnotowywane bezpośrednio w wypowiedziach użytkowników jednostki semantyczne, fonologiczne oraz, stanowiące ich powiązania, jednostki symboliczne, ewentualnie jednostki wyłaniające się z nich w wyniku działania procesów mentalnych poświadczonych w badaniach psychologicznych, takich jak kategoryzacja czy schematyzacja.

Należy jednocześnie zauważyć, że w odróżnieniu od podejścia generatywnego językoznawstwo kognitywne odrzuca możliwość wprowadzania do teorii języka semantycznie pustych symboli. Jak podkreśla Langacker:

Znaczenie jest istotą języka. Ignorowanie go poprzez skupienie wyłącznie na kwestiach formy prowadzi do drastycznego zubożenia [...] przedmiotu językoznawstwa, a w ostatecznym rozrachunku do zniekształcenia natury opisywanych zjawisk (Langacker, 1987, s. 12; przeł. Ł.W.).

O ile więc językoznawstwo generatywne charakteryzował związany z prymatem składni formalizm, o tyle cechą lingwistyki kognitywnej jest semantyzm, który przejawia się również założeniem, że wszelka, nawet nieznaczna zmiana formy językowej wypowiedzi pociąga za sobą zmianę w zakresie znaczenia.

Stanowisko współczesnej lingwistyki kognitywnej wobec natury znaczenia językowego cechuje subiektywizm. Zgodnie z tym ujęciem znaczenia wypowiedzi językowych nie da się ująć w kategoriach prawdy i fałszu ani wyrazić za pomocą sformalizowanego języka logiki, albowiem znaczenie nie stanowi odbicia stanu rzeczy w świecie rzeczywistym, lecz odzwierciedlenie sposobu, w jaki ten stan rzeczy pojmuje będący użytkownikiem języka człowiek (Tabakowska, 1995, s. 17). Ścisłej rzecz ujmując, znaczenie językowe utożsamiane jest z tzw. konceptualizacją, rozumianą jako doświadczenie mentalne (Langacker, 1987, s. 5). W przypadku odbiorcy komunikatu językowego (tj. słuchacza/czytelnika) chodzi o ogół doświadczenia mentalnego wywołanego czyjąś wypowiedzią ustną bądź lekturą danego fragmentu tekstu w jakimś konkretnym kontekście, natomiast w przypadku autora wypowiedzi

(tj. osoby mówiącej/piszącej) chodzi o doświadczenie mentalne prowadzące do użycia określonej formy językowej (wyrażenia, zdania itp.) na potrzeby komunikacji językowej. Co istotne, konceptualizacja ma charakter procesowy: nie chodzi tu jedynie o uruchamianie pojedynczych, statycznych pojęć będących elementem dotychczasowej wiedzy użytkownika języka, lecz o dynamiczny proces aktywacji treści pojęciowych w czasie przetwarzania, obejmujący również ich integrację i tworzenie nowych pojęć, a także przeżywane emocje i doznania zmysłowe. U podstaw rozumianej w ten sposób konceptualizacji leżą w ostatecznym rozrachunku określone wzorce aktywności neuronalnej i ma ona prawdopodobnie charakter wyobrażeniowy, a nie propozycjonalny (Langacker, 2009, s. 52–57).

Językoznawstwo kognitywne przyjmuje tzw. encyklopedyczną koncepcję znaczenia (Langacker, 1987, s. 154–166; Langacker, 2009, s. 64–65). Semantyka encyklopedyczna przeciwstawiana jest tradycyjnej koncepcji znaczenia językowego, zgodnie z którą treści semantyczne charakteryzowane są na wzór definicji słownikowej. Podejście słownikowe zakłada, że znaczenie jednostek leksykalnych (tj. słów) danego języka stanowi ograniczony zbiór informacji, wyrażany zwykle w postaci niewielkiego zbioru cech. Opiera się ono na założeniu, że możliwe jest precyzyjne wytyczenie granic między znaczeniem czysto językowym a wiedzą o świecie czy też wiedzą kontekstową, która w językoznawstwie tradycyjnym jest przedmiotem pragmatyki. Semantyka słownikowa nastęrcza jednak szeregu problemów w analizie znaczenia wyrażen językowych, chociażby z uwagi na fakt, iż poszczególne elementy znaczenia danego słowa mogą okazać się mniej lub bardziej istotne, czy nawet w ogóle irrelewantne, w różnych przypadkach jego użycia, a znaczenie złożonych jednostek leksykalnych nie zawsze da się ustalić jedynie na podstawie znaczeń ich elementów składowych. Ponadto określenie granic wiedzy językowej (semantycznej) jest w praktyce zadaniem niewykonalnym, ponieważ każde słowo stanowi w istocie punkt dostępu mentalnego do pewnego zbioru wiedzy, przy czym nie sposób przewidzieć wszystkich możliwych kierunków rozchodzenia się aktywacji treści pojęciowych w konceptualizacji związanej z użyciem danego słowa, w związku z czym należy uznać, że jest to zbiór wiedzy o charakterze otwartym. Wobec powyższego w językoznawstwie kognitywnym przyjmuje się podejście do analizy znaczenia,

zgodnie z którym treści pojęciowe powiązane z jednostkami leksykalnymi stanowią krystalizujący się ostatecznie w określonym kontekście, ustrukturyzowany zbiór szerokiej wiedzy o charakterze encyklopedycznym i nie da się wytyczyć ostrej granicy oddzielającej treści *stricte* językowe (semantyczne) od treści pozajęzykowych (pragmatycznych).

Zacieranie granic między tradycyjnie przeciwstawianymi sobie działami językoznawstwa i odpowiadającymi im elementami struktury języka – na wzór rozmytych granic w modelu kategoryzacji opartym na prototypie – jest we współczesnym językoznawstwie kognitywnym praktyką dość powszechną. Poza wspomnianym już rozmyciem podziału na semantykę i pragmatykę warto odnotować, kontrowersyjne na pozór, stanowisko Langackera względem jednostek leksykalnych i gramatycznych elementów języka. W przeciwieństwie do modelu generatywnego, w którym to zbiór jednostek leksykalnych (leksykon) i składnia posiadały zupełnie inny status (nieuporządkowany wykaz obiektów i ich cech *versus* system reguł) i należały do odrębnych komponentów języka, według Langackera granica między leksykonem a składnią jest płynna. Ściślej rzecz biorąc, słowa danego języka i jego elementy gramatyczne tworzą kontinuum jednostek symbolicznych, z których każda stanowi powiązanie dwóch biegunów: fonologicznego i semantycznego (Langacker, 2009, s. 20).

Pojęcie kontinuum odnosi się w tym przypadku do szczegółowości informacji dotyczących wymowy (w przypadku bieguna fonologicznego) oraz znaczenia (w przypadku bieguna semantycznego) jednostki symbolicznej. Jeden kraniec tego kontinuum zajmują konkretne jednostki leksykalne (np. rzeczownik *książka*, czasownik *czytać* czy przyimek *na*), których wymowa i znaczenie wyspecyfikowane są w sposób szczegółowy; na drugim zaś końcu znajdują się jednostki i struktury gramatyczne, w przypadku których zarówno znaczenie, jak i wymowa określone są w sposób schematyczny. Przykładem może być część mowy, taka jak rzeczownik, czasownik czy przyimek. Z jednej strony ich biegun fonologiczny zawiera bardzo niewiele informacji (jest schematyczny), jednak nie można powiedzieć, że jest zupełnie pusty – obejmuje choćby informacje na temat obowiązujących w danym języku wzorców i ograniczeń fonologicznych (np. każdy polski wyraz zawiera przynajmniej jedną samogłoskę; polskie przyimki proste są zwykle jednosylabowe itd.); z drugiej natomiast strony, schematycznie określone treści zawiera

również ich biegun semantyczny (w językoznawstwie kognitywnym – co również odróżnia je od ujęć strukturalistycznych i generatywnych – części mowy definiowane są na podstawie kryteriów pojęciowych/semantycznych, a nie formalnych), np. rzeczownik oznacza obszar w określonej przestrzeni pojęciowej, czasownik – proces, a przyimek – relację niezmienną w czasie.

W podobny sposób ogólnie scharakteryzowane znaczenie posiadają również struktury składniowe, na przykład konstrukcja z dwoma dopełnieniami leżąca u podstaw zdań typu: *Wysłałem ci e-mail, Firma wypłaciła pracownikom premię czy Studenci zadawali wykładowcy pytania*. Jej strukturę semantyczną można by opisać następująco: jedna rzecz działa na inną rzecz pewną siłą w taki sposób, że w rezultacie ta druga rzecz przechodzi w obszar kontroli rzeczy trzeciej (oczywiście przyjmując wspomniane powyżej specyficzne dla gramatyki kognitywnej rozumienie pojęcia rzeczy). Jeśli zaś chodzi o przypadki pośrednie, zajmujące środek wspomnianego kontinuum jednostek symbolicznych, obejmują one na przykład konstrukcje, których warstwa fonologiczna i znaczenie są scharakteryzowane w sposób mniej szczegółowy niż konkretne wyrażenia językowe, ale jednocześnie bardziej konkretny niż wysoce schematyczne elementy gramatyczne. Jako przykład niech posłuży struktura morfologiczna rzeczownika rodzaju męskiego z sufiksem deminutywnym {-ek}. Zwróćmy uwagę, że fonologiczny biegun takiej jednostki jest określony, z jednej strony, w sposób bardziej ogólny niż biegun fonologiczny konkretnych wyrazów stworzonych na podstawie tej struktury (np. *kotek, domek, papierosek*), z drugiej zaś w sposób bardziej konkretny w porównaniu z ogólną kategorią rzeczownika. To samo dotyczy jej bieguna semantycznego: jest on bardziej ogólny w porównaniu z konkretnymi przykładami takiego wyrazu, ale jednocześnie zawiera więcej informacji niż ogólna jednostka, jaką jest rzeczownik: oznacza rzecz małych rozmiarów lub rzecz postrzeganą jako miłą, ładną bądź w inny sposób pozytywnie ocenianą przez użytkownika języka.

Przedstawione powyżej proste przykłady ilustrują postulowane przez Langackera rozmycie umownych granic pomiędzy leksykonem, morfologią i składnią. Podkreślić należy pewną istotną implikację tezy o symbolicznej naturze całego języka: elementy tradycyjnie identyfikowane jako *stricte* gramatyczne niosą ze sobą znaczenie podobnie jak jednostki leksykalne, choć znaczenie to jest bardziej ogólne. Tym samym

językoznawstwo kognitywne odrzuca typowe dla generatywizmu traktowanie struktur gramatycznych w kategoriach wyłącznie formalnych oraz możliwość przypisywania tego samego znaczenia wypowiedziom odnoszącym się do tej samej sytuacji, lecz opartych na różnych konstrukcjach składniowych. W świetle założeń lingwistyki kognitywnej struktury gramatyczne są „usankcjonowanymi przez konwencję (i wobec tego zobiektywizowanymi) sposobami” konceptualizacji treści pojęciowych, czy też inaczej: „są to językowe obrazy świata widzianego oczyma «konwencjonalnego» obserwatora” (Tabakowska, 1995, s. 17).

Do najważniejszych kierunków badawczych językoznawstwa kognitywnego należą badania nad metaforą. Jak zauważa Taylor (2007, s. 587): „[c]hyba żadna inna książka nie przyczyniła się tak bardzo do spopularyzowania językoznawstwa kognitywnego, jak *Metafory w naszym życiu*” autorstwa Lakoffa i Johnsona (1988). Znamienny jest jednak również fakt, że była to pierwsza z kilku monografii (pierwotnie wydana w 1980 roku), które odegrały kluczową rolę w rozwoju podstaw teoretycznych współczesnego nurtu kognitywnego w językoznawstwie. Zaproponowana w niej koncepcja mierzyła się ze zjawiskiem redukowanym przez generatywistów do literackiego ornamentu i wobec tego konsekwentnie pomijanym, a jednocześnie czyniła to w sposób kognitywistyczny *par excellence*. Należy zaznaczyć, że od lat 80. XX wieku koncepcja ta uległa szeregowi modyfikacji i doczekała się wersji rozszerzonych, których nie będziemy tu omawiać, jednakże jej fundamentalne tezy pozostają w dużej mierze nadal aktualne.

Przed wszystkim Lakoff i Johnson zwracają uwagę na wszechobecność języka figuratywnego, w tym metafory: nie tylko w poezji, ale również w życiu codziennym, co zresztą sugeruje już sam tytuł ich przełomowej monografii. Fakt ten uzasadnia, a nawet nakazuje, włączenie metafory w zakres przedmiotowy językoznawstwa.

Kognitywistyczny sposób rozumienia tego zjawiska uwidacznia się już w nazwie koncepcji zaproponowanej przez Lakoffa i Johnsona: teoria metafory pojęciowej (*conceptual metaphor theory*). Istota kognitywnej teorii metafory sprowadza się właśnie do przypisywania jej charakteru mentalnego, pojęciowego, a nie ściśle językowego. Inaczej rzecz ujmując: metafora przedstawiana jest tutaj nie tylko jako sposób wyrażania się na płaszczyźnie językowej, lecz przede wszystkim jako sposób myślenia. Jak ujmują to autorzy: „[s]ystem pojęć, którymi się

zwykle posługujemy, by myśleć i działać, jest w swej istocie metaforyczny” (Lakoff & Johnson, 1988, s. 25). Metaforyczność polega tu mianowicie na pojmowaniu zbioru pojęć związanych z jednym obszarem ludzkiego doświadczenia w kategoriach zbioru pojęć odnoszących się do pewnego innego obszaru doświadczenia. Zgodnie z takim ujęciem podstawową funkcją metafory jest funkcja poznawcza, a nie stylistyczna.

Rozumiane w ten sposób zjawisko metafory wyrażane jest zwykle w postaci formuły: X TO Y (w językoznawstwie kognitywnym zapis dużymi literami symbolizuje obiekty pojęciowe), gdzie X oznacza zbiór pojęć odnoszących się do doświadczenia stanowiącego rzeczywisty przedmiot refleksji (określany mianem domeny docelowej), natomiast Y – zbiór pojęć, przez pryzmat którego pojmowane jest to doświadczenie (określany mianem domeny źródłowej). Mechanizmem leżącym u podstaw klasycznej wersji metafory pojęciowej jest odwzorowanie (w matematycznym sensie tego słowa) w domenie docelowej określonych elementów struktury domeny źródłowej. Warto jednocześnie zaznaczyć, że domena źródłowa, dostarczająca ram pojęciowych do konceptualizacji elementów domeny docelowej, obejmuje zwykle doświadczenie bardziej konkretne i uchwytnie zmysłami, podczas gdy domena docelowa odnosi się najczęściej do doświadczenia bardziej abstrakcyjnego i przez to trudniejszego do uchwycenia.

Przykładem metafory pojęciowej jest metafora MIŁOŚĆ TO PODRÓŻ, w której domenę źródłową stanowią elementy wiedzy dotyczące podróżowania, podczas gdy domeną docelową jest relacja romantyczna między dwiema osobami. Metafora ta polega na ustaleniu następujących odwzorowań pomiędzy elementami domeny źródłowej i docelowej: tworzące związek osoby z domeny docelowej odpowiadają podróżnikom z domeny źródłowej; wspólnemu celowi partnerów odpowiada miejsce docelowe podróży; wspólnym dotychczasowym doświadczeniom odpowiada pokonana droga; trudnościom w związku przyporządkowane są przeszkody na drodze itd. Jako mechanizm poznawczy metafora ta predysponuje nas do myślenia o związku dwojga osób w kategoriach podróży, co z kolei przekłada się na określony sposób językowego opisywania miłości: „Spójrz, *jak daleko zaszliśmy*. Jesteśmy *na rozdrożu*. Musimy się *rozstać*. Niech każde z nas *idzie swoją drogą*. Nie możemy teraz *zawrócić*” (Lakoff & Johnson, 1988, s. 68). Zgodnie z kognitywną koncepcją metafory konkretne figuratywne wyrażenia językowe, jak te

powyżej, są w pewnym sensie skutkiem ubocznym działania określonej metafory pojęciowej, która – jako sposób myślenia – jest wobec języka pierwotna. Warto zauważyć, że wymienione przykłady polskich wyrażen bazujących na metaforze MIŁOŚĆ TO PODRÓŻ nie należą do języka poetyckiego – reprezentują raczej potoczną mowę codzienną. Istnieje jednak zawsze możliwość twórczego rozwinięcia danej metafory pojęciowej poprzez stworzenie nowatorskiego wyrażenia, którego znaczenie pozostaje w zgodzie z jej sensem ogólnym. Na przykład stwierdzenie *Przelotny afekt wyniósł nas w stan nieważkości, z którego spadamy, świadomi niechybnego zderzenia z ziemią* mogłoby być prawdopodobnie uznane za poetycki opis miłosnej relacji bazujący na wymienionej metaforze pojęciowej. Przykład ten jednocześnie pokazuje, że granica między językiem codziennym a językiem literackim jest w gruncie rzeczy kwestią stopnia, a „wszystkie przejawy użycia języka czerpią z tych samych zasobów poznawczych” (Tabakowska, 2001, s. 46).

Do innych, często wymienianych przykładów metafor pojęciowych należą: ŻYCIE TO PODRÓŻ (*Miała w dzieciństwie pod górkę, Dotarł do kresu swojej wędrówki, Muszę trochę zwolnić i złapać oddech*), ARGUMENTOWANIE TO WOJNA (*czyjaś linia obrony, wytoczyć ciężkie działa, wytrącić komuś z ręki argumenty*), CZAS TO CENNY ZASÓB (*marnować czas, zyskać na czasie, zaoszczędzić trochę czasu*), SZCZĘŚLIWY TO W GÓRĘ (*podnieść kogoś na duchu, uszy do góry, być w dołku*), a także stanowiąca nasz sposób myślenia o komunikacji językowej tzw. metafora przewodu, będąca w istocie zespołem trzech metafor pojęciowych: ZNACZENIA TO PRZEDMIOTY, WYRAŻENIA JĘZYKOWE TO POJEMNIKI, KOMUNIKACJA TO PRZESYŁANIE (Lakoff & Johnson, 1988, s. 32). Wyrazem naszego codziennego pojmowania komunikacji językowej za pomocą metafory przewodu są takie wyrażenia jak: *przekazać komuś informację, puste słowa, słowa pełne pogardy, nie znajdować sensu w czyjejs wypowiedzi, odbierać czyjeś słowa w jakiś sposób* itd.

Podstawę dużej części metafor pojęciowych stanowi nasze doświadczenie fizyczne, z którego na bardzo wczesnym etapie rozwoju ontogenetycznego wyłaniają się powtarzalne wzorce doświadczenia przedpojęciowego, tzw. schematy wyobrazeniowe (*image schemata*) (Johnson, 1987). Tak na przykład jedną z przedpojęciowych struktur pojawiających się nieustannie w naszym codziennym doświadczeniu fizycznym jest

schemat wyobrażeniowy ŹRÓDŁO-ŚCIEŻKA-CEL, stanowiący immanentny element doświadczenia przemieszczania się z jednego miejsca w drugie. Innym przykładem jest schemat POJEMNIKA, wyłaniający się z fizycznych doświadczeń towarzyszących człowiekowi praktycznie od narodzin: już jako niemowlęta wkładani jesteśmy do różnego rodzaju pojemników (np. łóżeczko, wózek) stanowiących dla naszego ciała barierę, a później własnoręcznie umieszczamy w pojemnikach różne przedmioty (np. klocki w pudełku) bądź sami się w nich znajdujemy (np. w budynku czy w samochodzie). Jeszcze innym przykładem schematu wyobrażeniowego jest schemat GÓRA-DÓŁ, którego doświadczamy, na przykład, będąc podnoszonym przez matkę z kołyski i odkładanym z powrotem czy budując wieżę poprzez dokładanie kolejnych klocków, a następnie burząc ją i obserwując, jak w konsekwencji zmniejsza się jej wysokość. Tego rodzaju schematy wyobrażeniowe, wynikające ze specyfiki naszych ciał i naszego doświadczenia ze środowiskiem fizycznym, w którym żyjemy, nadają strukturę naszym pojęciom, a przez to wpływają również na sposób, w jaki mówimy. I tak, schemat ŹRÓDŁO-ŚCIEŻKA-CEL leży u podstaw naszego metaforycznego myślenia o życiu i miłości. Schemat GÓRA-DÓŁ stanowi bazę doświadczeniową dla wielu tzw. metafor orientacyjnych (Lakoff & Johnson, 1988, s. 36 i nast.), takich jak wspomniana już wyżej metafora SZCZĘŚLIWY TO W GÓRĘ czy WIĘCEJ TO W GÓRĘ (np. *wzrost/spadek cen, wysokie/niskie zarobki, wyż/niż demograficzny* itd.). Z kolei schemat POJEMNIKA staje się jednym z fundamentów rozumienia takich abstrakcyjnych domen naszego doświadczenia jak komunikacja językowa (SŁOWA TO POJEMNIKI NA ZNACZENIE) czy stany i sytuacje, które również pojmujemy na wzór pojemników (np. *być w trudnej sytuacji, w kryzysie, popaść w depresję, wpaść w euforię/szał* itd.).

Schematy wyobrażeniowe są tylko jednym z przejawów wpływu specyfiki ludzkiego ciała i jego fizycznej interakcji z otoczeniem na aparat poznawczy człowieka, a w konsekwencji również na język. Przekonanie o takim wpływie znajduje w językoznawstwie kognitywnym wyraz w postaci tezy o ucieleśnieniu umysłu i języka. Na znaczenie tej tezy dla kognitywistycznego nurtu w lingwistyce wskazuje tytuł monografii M. Johnsona (1987), który w dosłownym polskim przekładzie brzmi: „Ciało w umyśle: cielesne podstawy znaczenia, wyobraźni i rozumowania”. Uwzględnianie w teorii i badaniach nad językiem czynników

związanych z ucieleśnieniem ludzkiego umysłu jest więc programowo wpisane we współczesne kognitywne podejście do języka.

Refleksja systematyczna z wnioskami i rekomendacjami

Przedstawione powyżej – z konieczności skrótowo i wybiórczo – założenia, postulaty i kierunki badawcze współczesnej lingwistyki kognitywnej pokazują, w jaki sposób i dlaczego nauka zajmująca się językiem wpisuje się w szeroko zakrojony, program badań kognitywistycznych. Kluczowym powodem motywującym jej udział w kognitywistyce jest przekonanie, że zdolność używania języka bazuje na ogólnych właściwościach aparatu poznawczego człowieka, a struktura języka stanowi w wielu aspektach ich odzwierciedlenie. Zasadniczym zadaniem językoznawstwa, jakie się z takiego podejścia wyłania, jest zatem odkrywanie i opis zjawisk o charakterze mentalnym na podstawie autentycznych danych językowych. Sprostanie tak zakrojonemu celowi stanowi jednak poważne wyzwanie.

Po pierwsze, pomimo swego kognitywistycznego nachylenia i skupienia na języku jako zdolności poznawczej, lingwistyka kognitywna nie może abstrahować od tego, że odpowiadające za język procesy mentalne osadzone są nie tylko w ludzkim ciele, ale także w złożonym kontekście społeczno-kulturowym, co w konsekwencji znajduje swe odbicie w strukturze języka podobnie jak ucieleśnienie. Wartość teorii języka rozwijanej w ramach lingwistyki kognitywnej oraz jej wkładu w dorobek kognitywistyki zależą więc w istotny sposób od zakresu, w jakim uwzględnia ona czynniki społeczno-kulturowe. Zwrócenie większej uwagi na te kwestie oraz dalszy rozwój metodologicznych podstaw ich analizy w ramach lingwistyki kognitywnej stanowić będzie niewątpliwie pożądanym elementem humanistycznego zwrotu w dociekaniach kognitywistycznych (por. Bremer, 2023).

Drugie z zasadniczych wyzwań lingwistyki kognitywnej dotyczy jednego z najczęściej formułowanych wobec niej zarzutów: niemożliwości falsyfikacji tez formułowanych w odniesieniu do semantycznej struktury języka. Problem wynika tu z przyjęcia subiektywistycznej koncepcji znaczenia, utożsamiającej je z konceptualizacją. Wobec braku narzędzi

dających bezpośredni dostęp do świata mentalnego użytkowników języka podejście takie stwarza ryzyko bezwarunkowej akceptacji postmodernistycznego „relatywizmu metodologicznego” w analizie semantycznej, co „może sprawiać wrażenie sprowadzania badań naukowych do impresjonistycznych wprawek w interpretacji” (Tabakowska, 2001, s. 27). Sytuacji nie poprawia też w dalszym ciągu dość powszechnie stosowana praktyka formułowania tez semantycznych na podstawie danych pochodzących z introspekcji. Jednym ze sposobów podołania temu wyzwaniu jest zwiększenie możliwości replikacji badań semantycznych, na przykład poprzez – prężnie już zresztą rozwijane w ramach językoznawstwa kognitywnego – badania semantyczne oparte na metodach ilościowych i wykorzystaniu narzędzi statystycznych (m.in. technik eksploracyjnych oraz statystycznych metod weryfikacji hipotez). Odpowiedzią na obydwie z zarysowanych powyżej wyzwań jest również intensyfikacja interdyscyplinarnych kognitywnych badań nad językiem, w tym programów badawczych obejmujących wykorzystanie nowoczesnych technik badania procesów neurofizjologicznych.

Na koniec podkreślić należy znaczenie konfrontacji formułowanych przez lingwistów tez z relewantnymi ustaleniami, w tym najnowszymi, poczynionymi na gruncie pozostałych nauk kognitywnych. Jeżeli językoznawstwo kognitywne ma być naprawdę kognitywne, musi dochowywać wierności swojemu podstawowemu zobowiązaniu: zobowiązaniu do kognitywizmu.

BIBLIOGRAFIA

- Bremer, J. (2023). Humanistyka kognitywna. W: B. Bodzioch-Bryła (Red.), *Humanistyka współczesna*, (s. 71–88). Kraków: Wydawnictwo Naukowe Uniwersytetu Ignatianum w Krakowie.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1959). Review of B.F. Skinner's *Verbal Behavior*. *Language*, 35(1), 26–58.
- Chomsky, N. (1982 [1965]). *Zagadnienia teorii składni*. Przeł. I. Jakubczak. Wrocław: Zakład Narodowy im. Ossolińskich.
- Geeraerts, D., & Cuyckens, H. (2007). Introducing Cognitive Linguistics. W: D. Geeraerts, & H. Cuyckens (Red.), *The Oxford Handbook of Cognitive Linguistics* (s. 3–21). Oxford/New York: Oxford University Press.

- Heinz, A. (1983). *Dzieje językoznawstwa w zarysie*. Warszawa: PWN.
- Johnson, M. (1987). *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Reason*. Chicago/London: The University of Chicago Press.
- Lakoff, G. (1990). The Invariance Hypothesis: is abstract reason based on image-schemas?. *Cognitive Linguistics*, 1(1), 39–74.
- Lakoff, G. (2011 [1987]). *Kobiety, ogień i rzeczy niebezpieczne. Co kategorie mówią nam o umyśle*. Przeł. M. Buchta, A. Kotarba & A. Skucińska. Kraków: Universitas.
- Lakoff, G., & Johnson, M. (1988 [1980]). *Metafory w naszym życiu*. Przeł. T.P. Krzeszowski. Warszawa: Państwowy Instytut Wydawniczy.
- Langacker, R. (1987). *Foundations of Cognitive Grammar*. Vol. 1. *Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, R. (2009). *Gramatyka kognitywna. Wprowadzenie*. Przeł. E. Tabakowska & in. Kraków: Universitas.
- Taylor, J. (2007). *Gramatyka kognitywna*. Przeł. M. Buchta & Ł. Wiraszka. Kraków: Universitas.
- Taylor, J. (2001 [1989]). *Kategoryzacja w języku: prototypy w teorii językoznawczej*. Przeł. A. Skucińska. Kraków: Universitas.
- Tabakowska, E. (1995). *Gramatyka i obrazowanie. Wprowadzenie do językoznawstwa kognitywnego*. Kraków: Wydawnictwo Oddziału Polskiej Akademii Nauk.
- Tabakowska, E. (2001). *Językoznawstwo kognitywne a poetyka przekładu*. Przeł. A. Pokojska. Kraków: Universitas.

Mateusz Hohol

Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0003-0422-5488>

Nina Bażela

Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0009-0003-1147-1086>

<https://doi.org/10.35765/slowniki.436>

Poznanie matematyczne

Streszczenie

DEFINICJA POJĘCIA: Poznanie matematyczne to obszar nauki zajmującej się aktywnością umysłu związaną z matematyką, np. przetwarzaniem liczb i geometrii oraz biologicznymi podstawami kompetencji matematycznych człowieka. Poznanie matematyczne stanowi obszar badań interdyscyplinarnych, u których podstawy leżą psychologia i neuronauka.

ANALIZA HISTORYCZNA POJĘCIA: Badania nad poznaniem matematycznym poprzedza tradycja dociekań filozoficznych, dotyczących intuicji matematycznej. O ile według Platona, Kartezjusza, Kanta czy Poincarégo intuicja ta ma charakter wrodzony, o tyle według Helmholtza kształtowana jest ona na drodze doświadczenia. W prekursorskich badaniach nad rozwojem poznania matematycznego u dzieci Piaget wykorzystywał wciąż kategorię intuicji, przy czym wskazywał on, że wiedza o liczbach i przestrzeni budowana jest nie na drodze oglądu, lecz na podstawie aktywnego działania.

UJĘCIE PROBLEMOWE POJĘCIA: Poznanie matematyczne opiera się na systemach wiedzy rdzennej, które są stare ewolucyjnie, wykształcają się wcześniej w toku rozwoju osobniczego i są uniwersalne kulturowo. Umożliwiają one już we wczesnym wieku precyzyjną ocenę liczebności niewielkich zbiorów, szacowanie liczebności większych zbiorów, przetwarzanie kształtów oraz nawigację przestrzenną. Systemy te charakteryzują jednak ograniczenia, które przewyżczone są przy wsparciu języka w toku uczenia się liczebników oraz słownictwa przestrzennego. Badania wskazują ponadto, że w przejściu pomiędzy wiedzą zawartą w systemach

rdzennych a symboliczną matematyką istotną rolę odgrywają nawyki liczenia na palcach.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Badania nad poznaniem matematycznym dostarczają wiedzy o interakcjach natury i kultury, stanowiących jeden z podstawowych tematów nauk społecznych. Mają one również implikacje w obszarze edukacji, pomagając tworzyć wytyczne w zakresie wspierania osób z trudnościami szkolnymi.

Słowa kluczowe: geometria, liczby, asocjacje przestrzenno-numeryczne, wrodzoność, wiedza rdzenna

Definicja pojęcia

Z punktu widzenia kognitywistyki „poznanie matematyczne” rozumieć można na dwa sposoby – jako *explanans* i jako *explanandum* procesów poznawczych (Hohol, 2020). Pierwszy z nich odnosi się do wyjaśniania umysłu przy użyciu matematyki, w szczególności do modelowania różnorodnych procesów poznawczych przy wykorzystaniu narzędzi matematycznych. W drugim z nich aktywność poznawcza powiązana z wykorzystywaniem matematyki traktowana jest jako zjawisko wyjaśniane przy pomocy metod i teorii zorientowanej empirycznie kognitywistyki. Jednocześnie w drugim ujęciu poznanie matematyczne związane jest z obszarem badań interdyscyplinarnych, których rdzeń stanowią psychologia eksperymentalna i neuronauka poznawcza, uzupełniane przez dyscypliny, takie jak modelowanie obliczeniowe, antropologia i etologia poznawcza, a także badania edukacyjne.

W niniejszym artykule skupiamy się przede wszystkim na „poznaniu matematycznym” rozumianym jako *explanandum*, korzystając roboczo z ujęcia zaproponowanego przez Rafaela Núñeza i George’a Lakoffa (2005, s. 110). Zgodnie z tym ujęciem przedmiotem badań nad poznaniem matematycznym jest

pewien aspekt rzeczywistości psychologicznej, neurologicznej i edukacyjnej związanej z pewnymi zachowaniami, osiągnięciami i kompetencjami matematycznymi osoby, definiowanymi na poziomie jednostki lub na poziomie układu nerwowego jednostki, a matematyka jako taka pozostaje nietknięta.

Zakresem jest „to, co zazwyczaj rozumie się przez «matematykę», czyli elementarna arytmetyka, przetwarzanie liczb, a niekiedy podstawowa geometria lub elementarne myślenie algebraiczne”. Do przedmiotu badań należą również przestrzenne i ilościowe aspekty środowiska. Wreszcie badacze poznania matematycznego interesują sposoby uzasadniania i praktyki dowodowe stosowane w matematyce, a także przetwarzanie obiektów i struktur matematycznych przyswajanych na poziomie akademickim. Zakres badań nad poznaniem matematycznym jest szeroki i obejmuje badania nad różnymi populacjami – począwszy od niemowląt, aż po profesjonalnych matematyków. Metody badawcze „to głównie standardowe metody empiryczne stosowane w psychologicznych badaniach behawioralnych, badaniach nad zespołami

neuropsychologicznymi oraz obliczeniowym modelowaniu przetwarzania numerycznego” (Núñez & Lakoff, 2005, s. 110). Typowe pytania, z jakimi mierzą się badacze tak ujmowanego poznania matematycznego, są następujące: W jaki sposób umysł reprezentuje liczby? Czy istnieją wrodzone podstawy kompetencji numerycznych i geometrycznych człowieka, a jeśli tak, to jakie? Jakie struktury mózgowe zaangażowane są w elementarne przetwarzanie poznawcze liczb i własności geometrycznych? Czy poznawcze przetwarzanie struktur matematycznych przyswajanych na poziomie akademickim jest nadbudowane na elementarnym poznaniu numeryczno-geometrycznym?

Analiza historyczna pojęcia

Jednym z pierwszych i najważniejszych dzieł, które ukształtowały sposób myślenia o teoriach matematycznych, są bez wątpienia *Elementy* Euklidesa (IV–III w. p.n.e.), w których przedstawione zostały wzorce logicznego rozumowania i dedukcji matematycznej oraz podsumowana została wiedza matematyczna gromadzona na przestrzeni wieków przez matematyków greckich, którzy czerpali z kolei z dokonań Egipcjan i Babilończyków. Systematyzacja dokonana przez Euklidesa obejmowała geometrię na płaszczyźnie i stereometrię, a także algebrę geometryczną i teorię liczb, przy czym również ta ostatnia miała charakter geometryczny. Liczby ujmowane były jako odcinki, zaś mnożenie jako pomiar. Stąd też na przestrzeni wieków geometria pozostawała synonimem matematyki w ogóle, a dowody geometryczne uznawane były za wzór ścisłego myślenia. Jednym z terminów najczęściej pojawiających się w *Elementach* jest „konstrukcja”, o czym przekonać można się już od Księgi I, której pierwsze twierdzenie dotyczy konstrukcji trójkąta równobocznego na danej skończonej prostej. Choć termin „konstrukcja” sugeruje, że przed przeprowadzeniem rozumowania obiekty matematyczne nie istnieją, to na kartach dzieła Euklidesa nie znajdziemy tego rodzaju deklaracji filozoficznych.

Platon (IV w. p.n.e.) przekonany był, że przyswajanie wiedzy matematycznej w toku wieloletniego treningu możliwe jest dzięki wrodzonej zdolności poznawczej, która umożliwia uchwytywanie elementarnej wiedzy o punktach, odcinkach, kątach i relacjach między figurami.

Zdolność ta tradycyjnie określana jest jako „intuicja matematyczna” i uznać można ją za historyczną poprzedniczkę „poznania matematycznego” (Hohol, 2020). Intuicja, określana również mianem „wglądu”, powiązana jest z sięgającą co najmniej Platona metaforą „myślenia jako widzenia”. Tezę o intuicyjności elementarnej matematyki Platon formułuje m.in. w swoim dialogu *Menon*, w którym niewykształcony niewolnik wiedziony pytaniami zadawanymi przez Sokratesa okazuje się zdolny do przeprowadzenia rozumowania dedukcyjnego. Prowadzi go ono do wniosku, że kwadrat zbudowany na przekątnej danego kwadratu jest dwa razy większy (zob. Hohol, 2020). Według Platona w toku przeprowadzania konstrukcji geometrycznej powstaje reprezentacja obiektu matematycznego, „oglądana” przez umysł, ale istniejąca niezależnie od niego (zob. Hohol, 2020; Murawski, 2013). Podobnej tezy broniли kontynuatorzy myśli Platona, tacy jak Speuzyp z Aten (IV w. p.n.e.). Nie oznacza to jednak, że wszyscy filozofowie hellenistyczni byli platonikami. Jako kontrprzykład wymienić można choćby Arystotelesa, broniącego tezy o empirycznym charakterze wiedzy matematycznej, która powstaje w wyniku abstrakcji obiektów matematycznych z obiektów fizycznych, czy też Menaichmosa (IV w. p.n.e.) – związanego z Akademią Platońską matematyka, który twierdził, że konstrukcje geometryczne należy interpretować literalnie, tj. jako operacje, w wyniku których powstają obiekty matematyczne.

Intuicja matematyczna stała się jednym z klasycznych tematów filozofii nowożytnej. Kartezjusz (1596–1650) twierdził, że dotyczące wiecznych prawd twierdzenia matematyczne odkrywane są dzięki intuicji, którą nazywał „światłem naturalnym”. Co istotne, intuicja umożliwiająca rozpoznawanie jasności i wyraźności twierdzeń matematycznych jest według niego wrodzona i całkowicie niezależna od świadectw zmysłów, a także może być przybliżona poprzez metaforę „myślenie to widzenie” (zob. Hohol, 2020; Murawski, 2013). Dyskusje nad intuicją kontynuowano także w kolejnych epokach, w szczególności w odniesieniu do koncepcji Immanuela Kanta (1724–1804). Według tego ostatniego wszystkie twierdzenia matematyczne mają status tak zwanych sądów syntetycznych *a priori*, a więc takich, które z jednej strony rozszerzają naszą wiedzę (w przeciwieństwie do tautologii), a z drugiej są całkowicie niezależne od empirycznego uzasadnienia, przy czym ich prawdziwość rozpoznawana jest intuicyjnie (zob. Hohol, 2020; Murawski, 2013).

Możliwe jest to dzięki powiązaniu konstrukcji matematycznych z przestrzenią i czasem. Przestrzeń i czas rozumiane są tu jako „dane naoczne”, a zarazem jako wewnętrzne składniki umysłu, filtrujące wrażenia zmysłowe i organizujące poznanie. Wewnętrzna reprezentacja przestrzeni umożliwia podmiotowi konstrukcję geometrii, podczas gdy reprezentacja czasu leży u podstaw konstrukcji liczb, ujmowanych jako jednostki dołączane do siebie kolejno w czasie (zob. Hohol, 2020).

Do koncepcji intuicji matematycznej Kartezjusza i Kanta odwoływał się bezpośrednio Henri Poincaré (1854–1912), przy czym dokonał on daleko idących modyfikacji w jej rozumieniu (zob. Hohol, 2020). Poincaré określał intuicję matematyczną mianem „wrodzonej siły twórczej” ludzkiego umysłu, która umożliwia rozpoznawanie podstawowych relacji oraz twierdzeń matematycznych jako prawdziwych. Nowatorskim wkładem Poincarégo było natomiast zauważenie, że poznanie matematyczne angażuje zarówno świadome, jak i nieświadome przetwarzanie. Jego zdaniem twierdzenia matematyczne powstają początkowo na poziomie nieświadomym, ale muszą zostać uzupełnione przez świadomie kontrolowane rozumowanie. Wbrew Kantowi stwierdził on również, że aksjomaty, które służą jako punkt wyjścia dowodów geometrycznych, nie są sądami *a priori*, ale akceptowane są przez matematyków na mocy konwencji, w ustalaniu której niekiedy biorą oni pod uwagę fakty empiryczne, ale także kryteria takie jak wygoda i użyteczność. Doprowadziło to Poincarégo do uznania, że geometria definiowana przez zestaw aksjomatów zaprezentowanych przez Euklidesa nie jest jedyną możliwą, choć przyznawał on, że pozostaje ona najwygodniejsza. Wrodzona intuicja matematyczna gwarantuje więc zachowanie poprawności rozumowań, przy czym ostateczne wyniki zależne są od przyjętych na mocy konwencji założeń (zob. Hohol, 2020).

Jeszcze dalej w modyfikacji dotychczasowego rozumienia intuicji matematycznej poszedł Hermann von Helmholtz (1821–1894), zaprzeczając jej wrodzoności i skłaniając się ku radykalnemu empiryzmowi. Choć zgadzał się on z Kantem w tym, że doświadczenia i wrażenia zmysłowe są filtrowane przez umysł, odrzucił całkowicie transcendentálny punkt widzenia. Według Helmholtza zdolność poznawcza tradycyjnie określana jako intuicja matematyczna kształtowana jest na drodze akumulacji doświadczeń i wzmacniania powiązanych z nimi śladów pamięciowych. Oznacza to, że kształtowanie intuicji matematycznej można,

a nawet należy, „rozbić” na prostsze komponenty, a następnie wskazać, jak myślenie matematyczne kształtuje się w trakcie rozwoju osobniczego. Tezy Helmholtza na temat empirycznego pochodzenia intuicji matematycznej miały wciąż charakter filozoficzny, jednak inspirowane były jego badaniami psychologicznymi (należy wspomnieć, że był on jednym z pionierów psychologii eksperymentalnej) i położyły podwaliny pod badania empiryczne nad uczeniem się matematyki (zob. Hohol, 2020).

Termin „intuicja matematyczna” nie odszedł w zapomnienie wraz z powstaniem psychologii eksperymentalnej, choć jego znaczenie uległo kolejnej znaczącej zmianie. Istotną rolę w kształtowaniu się współczesnych badań empirycznych nad poznaniem matematycznym odegrał Jean Piaget (1896–1980). Uczony ten twierdził, że rozwój poznawczy przebiega wedle uporządkowanych faz, począwszy od fazy sensoryczno-motorycznej (od urodzenia do drugiego roku życia), w której dzieci zaczynają kształtować rozumienie stałości przedmiotów. Kolejna faza (trwająca do około szóstego roku życia) określana jest jako przedoperacyjna i charakteryzuje się brakiem możliwości przyjmowania perspektyw innych niż własna. Następnie (aż do dwunastego roku życia) dzieci przechodzą przez fazę operacji konkretnych, podczas której opanowują podstawy logiki i przyjmowania alternatywnych perspektyw, co umożliwi im rozumienie związków przyczynowo-skutkowych. Wreszcie w wieku dwunastu lat dzieci wkraczają w fazę operacji formalnych, która charakteryzuje się umiejętnością myślenia abstrakcyjnego.

Piaget poświęcił osobne prace genezie kompetencji numerycznych i geometrycznych, przy czym co do wszystkich pojęć matematycznych przyjmował on stanowisko konstruktywistyczne oraz zorientowane na działanie (zob. Semadeni, 2023). Jego zdaniem pojęcia matematyczne nie są wrodzone ani nie powstają nagle, ale są stopniowo konstruowane w umyśle dziecka na bazie doświadczeń. Nie kształtują się one w wyniku samego postrzegania świata fizycznego, ale eksploracji środowiska i aktywnej manipulacji przedmiotami. Jeśli chodzi o kompetencje arytmetyczne, Piaget wskazywał, że rozwijają się one równolegle ze zdolnościami logicznymi. Samo pojęcie liczby powstaje natomiast w wyniku abstrakcji (czyli pomijania cech fizycznych przedmiotów, takich jak kolor, wielkość czy kształt), bazującej na działaniach

dokonywanych na przedmiotach, takich jak np. przyporządkowywanie obiektów do zbiorów (zob. Szczygieł, 2017). Dla rozwoju pojęcia liczby kluczowe jest zinternalizowanie zasady stałości, a w szczególności zrozumienie, że zmiana położenia przedmiotów w zbiorach nie powoduje zmiany ich liczby. W typowych przeprowadzanych przez Piageta badaniach dziecku przedstawiano dwa równoliczne rzędy przedmiotów (np. kulek), znajdujących się w tej samej odległości, a następnie rozsuwano przedmioty w jednym z rzędów, co oznaczało zmianę układu przestrzennego, ale nie liczebności. Piaget zaobserwował, że nawet pięcioletnie dzieci twierdzą błędnie, że rozsuniecie prowadzi do zwiększenia liczby przedmiotów. Błąd ten ustępuje dopiero w szóstym lub siódmym roku życia, co oznacza pełne ukształtowanie pojęcia liczby. Późniejsze badania wykazały, że z zadaniami sprawdzającymi rozumienie zasady stałości obiektów radzą sobie również młodsze dzieci. Rozbieżność tę wyjaśniono m.in. nienaturalnością warunków badań przeprowadzanych przez Piageta (zob. Szczygieł, 2017). Co więcej, wykorzystanie innych paradygmatów badawczych umożliwiło badania nad obecnością reprezentacji liczb nawet u noworodków (zob. sekcję *Ujęcie problemowe pojęcia*).

Piaget przekonany był, że również pojęcia geometryczne nie wywodzą się po prostu z obserwacji świata. Przykładowo „równość” wywodzi się z internalizacji działania polegającego na wyrównywaniu dwóch przedmiotów, „prosta” ma u swych podstaw internalizację działania podążania za czymś, np. wzrokiem, bez zmiany kierunku, zaś „kąć” bierze się z internalizacji dwóch przecinających się ruchów np. rąk (zob. Hohol, 2020; Semadeni, 2023). Jeśli chodzi o poznawczą genezę geometrii, Piaget – jak już wspomniano – nie stronił od klasycznego pojęcia intuicji. Jak pisał on wspólnie z Bärbel Inhelder:

Intuicja przestrzeni nie jest czytaniem lub dostrzeganiem własności obiektów, ale od samego początku przeprowadzanym na nich działaniem. Ponieważ działanie wzbogaca i rozwija rzeczywistość fizyczną, zamiast jedynie ekstrahować z niej zbiór gotowych struktur, może ono ostatecznie wykroczyć poza ograniczenia fizyczne i stworzyć schematy operacyjne. Schematy te mogą zostać sformalizowane, tak by funkcjonowały w czysto abstrakcyjny i dedukcyjny sposób. Począwszy od podstawowej aktywności sensoryczno-motorycznej, aż po abstrakcyjne operacje, rozwój intuicji geometrycznej jest aktywnością w najpełniejszym znaczeniu tego słowa (Piaget & Inhelder, 1967, s. 449).

Koncepcja Piageta, choć w dużej mierze podważona przez późniejsze badania (zob. Spelke, 2022; Szczygieł, 2017), była oryginalna i historycznie ważna, stanowiła bowiem pierwszą próbę zdefiniowania etapów dojścia przez dziecko do pojęcia przestrzeni euklidesowej (zob. Hohol, 2020). Zdaniem Piageta rozwój indywidualny dokonuje się w porządku odwrotnym względem historycznego kształtowania się geometrii jako dziedziny matematyki. O ile geometria euklidesowa powstała już w starożytności, geometria rzutowa została opracowana w XIX w., zaś topologia powstawała na przełomie XIX i XX w., Piaget twierdził, że najbardziej pierwotnym pojęciem przestrzeni, kształtującym się w umyśle dziecka około trzeciego roku życia, jest „przestrzeń topologiczna” – trzeba jednak stwierdzić, że to, co Piaget rozumiał przez topologię, dalekie jest od ścisłego ujęcia matematycznego (zob. Hohol, 2020; Semadeni, 2023). Swoje wnioski sformułował m.in. na bazie eksperymentów, w których zadaniem dzieci było rozróżnianie kształtów np. na bazie dotyku i w których radziły sobie one dobrze, gdy jeden z przedmiotów był domknięty, a drugi nie. Najmłodsze dzieci nie radziły sobie natomiast z rozróżnianiem kwadratu od trójkąta ani nawet figur prostoliniowych od krzywoliniowych. Na tej podstawie Piaget stwierdził, że trzyletnie dzieci nie dysponują jeszcze pojęciami długości czy kąta. O ile w wieku około sześciu–siedmiu lat pewne pojęcia euklidesowe zdają się być już ukształtowane, wszystkie relacje geometryczne wciąż rozpatrywane są przez dziecko względem własnego punktu widzenia (co odpowiada wciąż egocentrycznej fazie przedoperacyjnej ogólnego rozwoju poznawczego). Na tej podstawie Piaget twierdził, że dziecko w tym wieku dysponuje pojęciem przestrzeni rzutowej. Wreszcie około dwunastego roku życia (co pokrywa się z wejściem w fazę operacji formalnych) dzieci rozwijają pojęcie przestrzeni euklidesowej, które nie odnosi się do konkretnego układu przedmiotów zajmujących postrzegane pozycje, ale raczej do organizacji samej przestrzeni jako „niewidzialnego pojemnika”, wewnątrz którego relacje są niezależne od aktualnego punktu widzenia (Piaget & Inhelder, 1967). Podobnie jak w przypadku rozwoju pojęcia liczby, teza o pierwszeństwie topologii jest współcześnie kwestionowana. Przykładowo wskazuje się na obserwację Piageta, że dzieci w pierwszej kolejności odróżniają figury otwarte od zamkniętych, jako wynikającą raczej z lepszej znajomości niektórych przedmiotów, a nie z najbardziej podstawowego charakteru topologii (zob. Hohol,

2020). Tym, co zdaje się natomiast aktualne, jest teza Piageta, zgodnie z którą kompetencje matematyczne nie wywodzą się z samej rejestracji świata za pomocą wzroku, ale z ruchu i działania w świecie.

Ujęcie problemowe pojęcia

Interdyscyplinarna nauka o poznaniu, znana dziś jako kognitywistyka, powstała na przełomie lat 50. i 60. XX w., obejmując początkowo przede wszystkim psychologię poznawczą, lingwistykę i informatykę, do których dołączały stopniowo neuronauka, antropologia i filozofia. Rozumowania matematyczne i logiczne stanowiły jedno z podstawowych zainteresowań wielu pionierów kognitywistyki, takich jak Allen Newell i Herbert Simon. Świadczą o tym współtworzone przez nich programy, określane jako pierwsze programy sztucznej inteligencji: Teoretyk Logiczny (*Logic Theorist*) i Uogólniony Rozwiązywacz Problemów (*General Problem Solver*), które zdolne były m.in. do radzenia sobie z dowodzeniem twierdzeń ze słynnego traktatu matematycznego *Principia mathematica* Bertranda Russella i Alfreda North Whiteheda. Inne programy pochodzące z tego okresu dowodziły natomiast z powodzeniem twierdzeń geometrii euklidesowej. Ponieważ krytycy wskazywali, że programy komputerowe mogą osiągać analogiczne wyniki matematyczne w zupełnie inny sposób niż ludzie, aby wyeliminować ten problem projektanci i programiści, zaczęli stopniowo wykorzystywać dane psychologiczne (pierwotnie protokoły werbalne, później dane dotyczące czasów reakcji). Stąd też programy geometryczne symulować zaczęły interakcje, w jakie ludzie wchodzi z reprezentacjami zewnętrznymi, np. diagramami (zob. Hohol, 2020).

Dowodzenie twierdzeń, będące zaawansowaną czynnością angażującą wiele procesów poznawczych, nie stało się jednak głównym tematem badań nad poznaniem matematycznym. Zamiast tego koncentrowano się na zagadnieniach bardziej elementarnych, takich jak sposób reprezentowania liczb przez umysł, który – jak się okazało – ma wiele wspólnego z analogowym przetwarzaniem wielkości fizycznych w ogóle. Wyniki licznych badań wskazują, że porównując dwa postrzegane przedmioty fizyczne, zarówno ludzie, jak i zwierzęta radzą sobie lepiej (osiągają większą poprawność i krótsze czasy reakcji), gdy

przedmioty znacząco różnią się pod względem wielkości, niż gdy różnica w ich wielkości jest nieduża. Gdy zaś badani porównują pary przedmiotów o stałej różnicy wielkości, łatwiej przychodzi im porównywanie par małych przedmiotów niż par większych obiektów (zob. Butterworth, 2022). Pierwsze z tych zjawisk określa się mianem efektu dystansu, zaś drugie – efektu rozmiaru. Co istotne, efekty te zachodzą, gdy przetwarzane są różne wielkości fizyczne, także takie jak długość odcinka czy jasność obiektu.

W 1967 r. Moyer i Landauer opisali podobny efekt dystansu dla liczb (zob. Brożek & Hohol, 2017; Dehaene, 2011). W zadaniu polegającym na wskazywaniu, która z dwóch liczb (przedstawianych za pomocą symboli arabskich) jest większa, ludzie reagują szybciej i popełniają mniej błędów, gdy porównywane liczby dzieli większy dystans (np. 5 i 9), niż gdy dystans ten jest mniejszy (np. 4 i 5). Odkryto również efekt rozmiaru dla liczb: gdy odległości między porównywanymi liczbami są takie same (np. różnica pomiędzy 3 i 5 oraz 5 i 7 wynosi 2), osoby badane udzielają odpowiedzi szybciej i popełniają mniej błędów w przypadku mniejszych liczb. Na bazie odkrycia tych dwóch efektów zaproponowano koncepcję umysłowej osi liczbowej (ang. *mental number line*). Efekt dystansu wskazuje na to, że na poziomie automatycznego i nieświadomego przetwarzania poznawczego reprezentacje liczb są ustrukturyzowane przestrzennie; porównanie dwóch liczb wymaga poznawczego „przybliżenia” określonego fragmentu osi, co łatwiejsze jest w przypadku liczb, które są od siebie bardziej oddalone. Z kolei efekt rozmiaru wskazuje na to, że inaczej niż w przypadku osi liczbowej znanej ze szkoły, umysłowa oś liczbową jest skompresowana logarytmicznie, tak że wraz ze wzrostem wielkości liczbowych maleją odległości między nimi; porównywanie mniejszych liczb jest zatem łatwiejsze ze względu na ich większą rozdzielczość (zob. Cipora & Nęcka, 2012). Niektóre badania sugerują, że w trakcie rozwoju osobniczego, a konkretnie wraz z nabywaniem szkolnych kompetencji matematycznych, stopień logarytmicznej kompresji osi liczbowej ulega zmniejszeniu, co wiąże się z redukcją siły efektów dystansu i rozmiaru. Z drugiej strony większość badaczy zgadza się dziś co do tego, że poznawcza reprezentacja wielkości, w tym wielkości liczbowej, może być aktywowana niezależnie od przestrzeni, co sprawia, że efekty dystansu i rozmiaru nie mogą być w łatwy sposób interpretowane na korzyść istnienia umysłowej osi liczbowej.

Teoretyczny konstrukt umysłowej osi liczbowej zyskał jednak poważne wsparcie w zaobserwowanym po raz pierwszy przez Stanisława Dehaene'a i współpracowników w 1993 r. i zreplikowanym odtąd kilkaset razy efekcie SNARC (Dehaene, 2011; zob. także Cipora & Nęcka, 2012; Brożek & Hohol, 2017). Efekt SNARC (*Spatial-Numerical Association of Response Codes*), czyli zależność przestrzenna między liczbą a rodzajem odpowiedzi, polega na tym, że w eksperymentach, w których zadanie osób badanych polega na podejmowaniu decyzji dotyczących właściwości prezentowanych kolejno liczb (np. wskazywaniu, czy liczba jest parzysta czy nieparzysta), obserwuje się krótsze czasy reakcji lewą ręką dla względnie małych liczb i prawą ręką dla względnie dużych liczb. Względnie, ponieważ efekt ten daje się zaobserwować nawet, gdy zbiór bodźców testowych obejmuje liczby od 1 do 9 (przy czym ze względów metodologicznych pomija się liczbę 5). Wzorzec ten sugeruje, że analogicznie do szkolnej osi liczbowej umysłowa reprezentacja liczb ustrukturyzowana jest kierunkowo – mniejsze liczby znajdują się bardziej po lewej, zaś większe po prawej stronie. Obecnie toczy się wciąż debata nad czynnikami wpływającymi na siłę efektu SNARC, a nawet sam kierunek (zob. Butterworth, 2022). Z jednej strony podkreśla się rolę czynników kulturowych w ukierunkowaniu umysłowej osi liczbowej. Standardowy efekt SNARC obserwuje się w kulturach, w których pisanie i czytanie przebiega od lewej do prawej, zaś odwrócony efekt SNARC w kulturach, gdzie pisze i czyta się od prawej do lewej. Co więcej, osoby, które mając policzyć na palcach od 1 do 10, zaczynają od lewej ręki, przejawiają silniejszy efekt niż te, które zaczynają od prawej. Z drugiej strony na podstawie wyników badań przeprowadzanych na zwierzętach wskazuje się, że samo wiązanie liczebności z przestrzenią ma głębokie korzenie ewolucyjne.

Efekty dystansu i rozmiaru wskazują na to, że liczby są reprezentowane poznawczo w sposób analogowy, czyli tak jak wielkości fizyczne. Z kolei efekt SNARC interpretuje się standardowo jako argument na rzecz powiązania umysłowych reprezentacji liczb i przestrzeni. Wnioski te wspierane są nie tylko przez wyniki licznych badań behawioralnych, ale również neuronaukowych, przeprowadzonych z wykorzystaniem funkcjonalnego obrazowania rezonansem magnetycznym. Przykładowo wskazuje się, że obszar bruzdy śródcieniowej mózgu wykazuje podobne reakcje, gdy osoby badane przetwarzają zarówno

liczby symboliczne, jak i odcinki. Na bazie tych oraz innych wyników Vincent Walsh sformułował teorię wielkości znaną jako ATOM (*A Theory of Magnitude*). Zakłada ona, że w korze płata ciemieniowego istnieje wspólna część mechanizmu metrycznego, zaangażowanego w przetwarzanie czasu, przestrzeni i liczb (zob. Butterworth, 2022). Warto pamiętać, że nie oznacza to, że wszystkie te wielkości przetwarzane są dokładnie tak samo czy że ich mechanizmy nie angażują dodatkowych komponentów. Co ciekawe, dowody świadczące o roli płata ciemieniowego w przetwarzaniu liczb oraz przestrzeni pojawiły się na długo przed tym, jak zaczęto stosować techniki neuroobrazowania. Jeszcze w pierwszej połowie XX w. austriacki neurolog Josef Gerstmann opisał zespół objawów (znany dziś jako zespół Gerstmannna) towarzyszących uszkodzeniu zakrętu kąтового, na które składają się trudności w przetwarzaniu liczb, odróżnianiu stron i pisaniu oraz agnozja palców (polegająca np. na zaburzeniach w rozpoznawaniu, który z palców pacjenta jest aktualnie stymulowany).

Współczesne wysiłki badawcze zmierzają do zrozumienia mechanizmu przetwarzania liczebności nie tylko na poziomie całych struktur mózgowych, ale także z uwzględnieniem poziomu pojedynczych neuronów. Badając makaki, Andres Nieder (2019) odnalazł tzw. numerony (ang. *numerical neurons*), znajdujące się w okolicy bruzdy śródcieniowej, czyli w obszarze wyspecjalizowanym w przetwarzaniu wielkości, oraz w bocznej korze przedczołowej, czyli strukturze zaangażowanej w ogólne procesy poznawcze. Ich nazwa bierze się stąd, że komórki te reagują selektywnie na małe liczebności niesymboliczne – dany neuron reaguje najsilniej, gdy makak postrzega jedną kropkę, inny neuron, gdy zwierzę postrzega dwie kropki i tak dalej, aż do czterech. Cztery (plus minus jeden element) stanowi barierę tzw. subitacji (ang. *subitizing*), czyli bezwysiłkowej, szybkiej i precyzyjnej oceny liczby elementów. Subitację należy odróżnić od szacowania, które polega na przybliżonym określeniu ilości elementów w zbiorze o liczebności większej niż cztery (plus minus jeden) oraz od liczenia. O ile subitacja i szacowanie są niezależne od reprezentacji symbolicznych, takich jak cyfry arabskie czy liczebniki, o tyle nauczony się liczenia, jesteśmy w stanie określić precyzyjnie liczebności zbiorów, które przekraczają – i to nawet znacznie – zakres subitacji. Choć przeprowadzono wiele badań wskazujących na to, że przedstawiciele niektórych gatunków zwierząt można

nauczyć liczenia, panuje konsensus, że zdolność ta jest częścią ekologii poznawczej człowieka (zob. Butterworth, 2022).

W kontekście dyskusji omówionych w sekcji *Analiza historyczna pojęcia* nasuwają się dwa pytania: czy w świetle obecnej wiedzy kognitywistycznej elementarne zdolności numeryczne, czyli subitację i szacowanie, można uznać za podstawę, na której w trakcie ontogenezy nadbudowywane są liczenie, arytmetyka i inne zdolności matematyczne? I czy te elementarne zdolności są wrodzone? Innymi słowy: czy rację mieli filozofowie szukający wrodzonych podstaw intuicji matematycznej, czy raczej Hermann von Helmholtz, który podważał istnienie czegokolwiek wrodzonego, prócz samych mechanizmów uczenia się? Choć wokół tych kwestii wciąż toczą się dyskusje (zob. Butterworth, 2022; Haman & Gut, 2016; Nieder, 2019), wielu badaczy poznania matematycznego skłonnych jest opowiedzieć się za wrodzonością elementarnego poznania numerycznego oraz tym, że stanowi ono podstawę dla uczenia się matematyki.

Jeśli chodzi o wrodzoność elementarnych kompetencji numerycznych, ludzkie niemowlęta przejawiają zdolność do odróżniania zbiorów różniących się liczbą przedmiotów, nie potrzebując do tego żadnego treningu. Wykazały to badania wykorzystujące tzw. paradygmat habituacji. W przypadku gdy dzieci w wieku około sześciu miesięcy obserwowały serię obrazków z jednakową liczbą czarnych punktów, które różniły się wielkością lub rozmieszczeniem, następowało skrócenie okresu ich zainteresowania danym bodźcem, co wskazuje na pojawienie się zjawiska habituacji. Natomiast gdy zaprezentowano im obrazy z odmienną liczbą punktów, bez względu na ich układ przestrzenny, odzyskiwały zainteresowanie obrazem, co objawiało się wydłużonym czasem obserwacji bodźca (zob. Brożek & Hohol, 2017). Inne badania wykazały natomiast, że już niemowlęta dysponują intuicją równoliczności. Przykładowo słysząc ciąg czterech sylab, noworodki spoglądały dłużej na zbiór czterech kropek niż na zbiór złożony z dwunastu kropek. Co więcej, badania przeprowadzone na licznych gatunkach zwierząt – począwszy od naczelnych różnych od człowieka, przez gryzonie, ptaki, a nawet ryby – wskazują, że zdolność do postrzegania liczebności jest powszechna w przyrodzie, co najmniej wśród kręgowców (zob. Butterworth, 2022; Nieder, 2019).

Zgodnie z wpływową i wciąż rozwijaną teorią systemów rdzennych autorstwa Elizabeth Spelke (2022), nasze zdolności poznawcze

opierają się na tzw. systemach wiedzy rdzennej (ang. *core knowledge systems*). Systemy te, jak pisze Spelke (2022),

przetwarzają ograniczoną dziedzinę obiektów i uchwytyją jedynie podzbiór własności, których dostarczają nasze zmysły. Systemy te są zamknięte, aktywowane automatycznie i nieświadome (...). Systemy wiedzy rdzennej pojawiają się na wczesnych etapach ontogenezy i funkcjonują przez resztę życia. Są one obecne u ludzi w każdym wieku i w każdej kulturze. Kierują one naszym myśleniem i uczeniem się zarówno w okresie niemowlęcym, jak i później. Co więcej, te same możliwości i ograniczenia systemów znaleźć można u szerokiej gamy zwierząt i zależą one od homologicznych systemów i procesów mózgowych, co dostarcza dowodów na to, że systemy rdzenne pojawiły się głęboko w naszej ewolucyjnej przeszłości (s. 190; zob. także Haman & Gut, 2016).

Wedle teoretyków systemów rdzennych za subitację i szacowanie odpowiadają dwa zlokalizowane w płacie ciemieniowym rozłączne systemy, określane kolejno jako system śledzenia obiektów oraz system przybliżania liczebności. Razem służą nam one za „zmysł liczebności” (ang. *the number sense*), który stanowi warunek konieczny dla opanowania liczb wyrażanych symbolicznie i wykonywania na nich działań (zob. Dehaene, 2011).

Jak wskazano powyżej, istnieją argumenty pozwalające twierdzić, że reprezentacje liczb mają charakter przestrzenny. Skąd biorą się jednak nasze intuicje przestrzenne i czy są one – jak twierdził Kant – podstawą geometrii euklidesowej? Według Spelke (2022) podobnie jak w przypadku „zmysłu liczebności”, nasze zdolności geometryczne ugruntowane są w dwóch ewolucyjnie starych, wczesnych ontogenetycznie i uniwersalnych kulturowo systemach wiedzy rdzennej, które określa się jako system geometrii przestrzennej i system geometrii obiektowej (Hohol, 2020). Pierwszy z nich zlokalizowany jest w formacji hipokampa i przetwarza odległości i kierunki, umożliwiając nawigację w środowisku. Drugi zlokalizowany jest natomiast w bocznych strukturach płata potylicznego i przetwarza kąty oraz długości, co umożliwia rozpoznawanie dwuwymiarowych kształtów i trójwymiarowych obiektów. Jak zauważa Spelke (2022), wszystkie systemy rdzenne cechują się ograniczeniami. System śledzenia obiektów dostarcza dokładnej wiedzy o liczebności, ale w zakresie zaledwie do około czterech przedmiotów. Z kolei system liczb przybliżonych pozwala operować na większych liczebnościach, lecz kosztem precyzji. System geometrii przestrzennej jest czuły na odległość

i kierunek, ale nie na kąty. Wreszcie system geometrii obiektowej przetwarza długości i kąty, ale nie kierunki. Jak jest to zatem możliwe, że reprezentacje poznawcze generowane przez tak ograniczone systemy łączone są w wolny od ograniczeń system pojęć matematycznych?

Choć wiele szczegółów pozostaje wciąż nieznanymi, według Spelke (2022) łączenie dokładnych reprezentacji małych liczb z przybliżonymi reprezentacjami większych liczb możliwe jest dzięki opanowaniu liczebników. Już w bardzo młodym wieku dzieci zaczynają posługiwać się podstawowymi liczebnikami, jednak do ukończenia drugiego roku życia nie są w stanie nadawać im właściwych znaczeń. Znacząca zmiana zachodzi w trzecim roku życia, kiedy dzieci zaczynają pojmować, że liczebnik „jeden” reprezentuje abstrakcyjną wartość 1. Proces ten jest kontynuowany dla następnych liczb i odpowiadających im liczebników do liczby 4 w czwartym roku życia. Od tego momentu zdolność do korzystania z liczebników rozwija się bardzo dynamicznie, co idzie w parze z nauką rozwiązywania bardziej skomplikowanych zadań arytmetycznych. Dzieci formułują uniwersalną zasadę pozwalającą na przyporządkowanie każdej liczbie odpowiadającego jej liczebnika, co sprzyja zrozumieniu zarówno aspektu kardynalnego, jak i porządkowego liczby. Język funkcjonuje więc jako rusztowanie, umożliwiające przełamanie ograniczeń rdzennych systemów numerycznych. Badania wskazują ponadto, że w przejściu pomiędzy wiedzą zawartą w systemach rdzennych a symboliczną matematyką istotną rolę odgrywają nawyki liczenia na palcach. Palce są obiektem konkretnym i łatwym do manipulowania, a zarazem nadają się do reprezentowania liczebności różnych przedmiotów. Co więcej, ze względu na naturalne ustrukturyzowanie, nawyki liczenia na palcach sprzyjają opanowaniu porządkowego aspektu liczby, podczas gdy systemy wiedzy rdzennej kodują jedynie aspekt kardynalny (zob. Szczygieł et al., 2015).

Badania wskazują, że w dziedzinie geometrii podobną do liczebników rolę odgrywa uczenie się słownictwa przestrzennego (zob. Hohol, 2020). Proces ten wspierany jest również przez rosnące doświadczenie dzieci w zakresie posługiwania się szkicami przypominającymi mapy, za sprawą których dzieci uczą się przedstawiać układy przestrzenne, reprezentowane pierwotnie przez system geometrii obiektowej (kierunek, odległość), jako dwuwymiarowe struktury, których kluczowymi właściwościami geometrycznymi są kąty i długości.

Refleksja systematyczna z wnioskami i rekomendacjami

Współczesne badania nad poznaniem matematycznym postrzegać można jako kontynuację dociekań filozoficznych nad naturą intuicji matematycznej. Świadczy o tym choćby to, że Stanislas Dehaene i Elizabeth Brannon (2010) nie wahają się nazywać jednego z głównych nurtów badań „Kantowskim programem badawczym”. W programie tym podkreśla się ciągłość między biopsychologicznymi podstawami poznania matematycznego a tym, w jaki sposób nasze umysły przetwarzają czas i przestrzeń, przy czym Kantowskie „filtry” zastępowane są rdzennymi systemami poznawczymi o długim rodowodzie ewolucyjnym. Dehaene i Brannon nie wahają się nawet stwierdzić, że gdyby Kant żył współcześnie, zostałby on nie filozofem, lecz neurokognitywistą (s. 519). W porównaniu z wcześniejszymi dociekaniem filozoficznymi współczesne badania nad poznaniem matematycznym umożliwiają wyjście poza sferę racjonalnych spekulacji i budowanie teorii możliwych do testowania w eksperymentach wykorzystujących metody, które sprawdziły się dobrze także w odniesieniu do innych sfer działania umysłu. Z drugiej jednak strony pochopte byłoby stwierdzenie, że empiryczne badania nad poznaniem matematycznym w prosty sposób zastępują filozofię matematyki. O ile trudno wyobrazić sobie dziś budowanie wiarygodnych teorii filozoficznych na temat wiedzy matematycznej, które byłyby jawnie sprzeczne z aktualnym rozumieniem mechanizmów neuropoznawczych, równie trudno wyobrazić sobie, by neurokognitywiści rozwiązali problem statusu ontologicznego obiektów i struktur matematycznych. Innymi słowy współczesne badania nad poznaniem matematycznym nie mogą rozstrzygnąć wspomnianego wcześniej sporu pomiędzy Speuzypem i Menaichmosem.

Innym istotnym wątkiem jest to, że empiryczne badania nad poznaniem matematycznym rzucają nowe światło na szereg problemów nurtujących w zasadzie wszystkie nauki społeczne. Kluczowy z nich wiąże się z pytaniem „natura czy kultura”. Z jednej strony podstawy poznania matematycznego mają charakter pozajęzykowy i związane są bezpośrednio ze zdolnościami wielu zwierząt w zakresie przetwarzania liczebności i umiejętnościami przestrzennymi, przez co wielu badaczy nie waha się uznawać ich za wrodzone (zob. Butterworth, 2022;

Haman & Gut, 2016). Z drugiej strony wrodzona forma poznania matematycznego, tj. wiedza rdzenna, jest bardzo ograniczona, by wspomnieć tylko o wąskim zakresie subitacji. Przekroczenie ograniczeń wiedzy rdzennej jest możliwe dzięki specyficznemu ludzkiemu wytworowi kulturowemu, czyli językowi, o czym świadczy choćby to, że użytkownicy języków, charakteryzujących się wąskimi systemami liczebników – należy do nich choćby lud Mundurucu – wykonują mniej precyzyjne obliczenia poza zakresem subitacji (zob. Brożek & Hohol, 2017; Butterworth, 2022). Klasyczne pytanie „natura czy kultura” zdaje się więc o tyle źle postawione, że bez wrodzonych systemów rdzennych zapewne w ogóle nie byłoby praktyk matematycznych przyswajanych przez nas w toku edukacji. Praktyki te zanurzone są jednak głęboko w kulturze, której wytwory działają jako „narzędzia poznawcze”, modyfikujące istotnie wiedzę rdzenną.

Z badaniami nad poznaniem matematycznym wiąże się również szereg ważkich implikacji praktycznych w zakresie edukacji. Specjaliści z obszaru poznania matematycznego rekomendują przede wszystkim, by matematyczne abstrakcje wprowadzane były w szkole poprzez gruntowanie ich w doświadczeniu przestrzennym, które wspólne jest wszystkim ludziom ze względu na strukturę naszych ciał. Szczególnie ważne jest w tym zakresie wspieranie tworzenia asocjacji przestrzenno-numerycznych. Teza, zgodnie z którą „im więcej przestrzeni, tym lepsza matematyka”, nie jest jednak uniwersalnie prawdziwa. O ile w przypadku dzieci silniejsze asocjacje przestrzenno-numeryczne (wyrażane np. przez umówiony wyżej silniejszy efekt SNARC) zazwyczaj wiążą się lepszymi kompetencjami w zakresie szkolnej matematyki, o tyle relacja ta może być nawet odwrotna w wieku dorosłym. Jest to szczególnie wyraźne w przypadku profesjonalnych matematyków, którzy w ogóle nie przejawiają efektu SNARC, co świadczyć może o tym, że długotrwały trening prowadzi do budowania bardziej abstrakcyjnych, czy też mniej ucieleśnionych, reprezentacji liczb (Cipora et al., 2016).

Co więcej, badania nad „zmysłem numerycznym” są cenne dla zrozumienia natury dyskalkulii rozwojowej, która utrudnia, a w skrajnych przypadkach nawet uniemożliwia przyswojenie szkolnej wiedzy matematycznej, stanowiąc barierę dla wkroczenia na liczne ścieżki zawodowe. Jak wskazują cyklicznie publikowane raporty edukacyjne, trudności związane z uczeniem się matematyki mają daleko idące konsekwencje

dla dobrobytu całych społeczeństw (zob. Butterworth, 2022, s. 30–33). Wobec tego rekomenduje się podjęcie interwencji psychologicznych, mających pomóc jak najmłodszym osobom z diagnozą dyskalkulii lub zagrożonych dyskalkulią, tak by stymulować możliwie najpełniejszy rozwój zmysłu numerycznego, co stanowi warunek konieczny dla przyswajania szkolnej wiedzy matematycznej (zob. Dehaene, 2011; Butterworth, 2022).

Wreszcie, badania z zakresu poznania matematycznego pokazują, że przetwarzanie wiedzy matematycznej, uważanej za najbardziej abstrakcyjny przedmiot ludzkiego poznania, ma u swych podstaw mechanizmy, które każda osoba – niezależnie od ścieżki edukacyjnej i zawodowej – wykorzystuje w codziennym życiu.

BIBLIOGRAFIA

- Brożek, B., & Hohol, M. (2017). *Umysł matematyczny*. Kraków: Copernicus Center Press.
- Butterworth, B. (2022). *Can fish count? What animals reveal about our uniquely mathematical mind*. London: Quercus.
- Cipora, K., Hohol, M., Nuerk, H.-C., Willmes, K., Brożek, B., Kucharzyk, B., & Nęcka, E. (2016). Professional mathematicians differ from controls in their spatial-numerical associations. *Psychological Research*, 80(4), 710–726. DOI: 10.1007/s00426-015-0677-6.
- Cipora, K., & Nęcka, E. (2012). Kontinua a przestrzeń – przegląd badań nad przestrzennym komponentem poznawczej reprezentacji wielkości i nasilenia. *Psychologia – Etologia – Genetyka*, 26, 7–21.
- Dehaene, S. (2011). *The number sense*. New York: Oxford University Press.
- Dehaene, S., & Brannon, E.M. (2010). Space, time, and number: A Kantian research program. *Trends in Cognitive Sciences*, 14(12), 517–519. DOI: 10.1016/j.tics.2010.09.009.
- Haman, M., & Gut, A. (2016). Wiedza wrodzona. W: J. Bremer (Red.), *Przewodnik po kognitywistyce* (s. 681–712). Kraków: Wydawnictwo WAM.
- Hohol, M. (2020). *Foundations of geometric cognition*. London: Routledge.
- Murawski, R. (2013). *Filozofia matematyki: Zarys dziejów*. Poznań: Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza.

- Nieder, A. (2019). *A brain for numbers: The biology of the number instinct*. Cambridge, MA: The MIT Press.
- Núñez, R., & Lakoff, G. (2005). The Cognitive Foundations of Mathematics: The Role of Conceptual Metaphor. W: J.I.D. Campbell (Red.), *Handbook of Mathematical Cognition* (s. 109–124). Psychology Press. DOI: 10.4324/9780203998045.109.
- Piaget, J., & Inhelder, B. (1967). *The child's conception of space*. New York: W.W. Norton & Company.
- Semadeni, Z. (2023). *Różne oblicza matematyki*. Toruń: Wydawnictwo Naukowe UMK.
- Spelke, E.S. (2022). *What babies know: Core knowledge and composition. Volume 1*. Oxford University Press. DOI: 10.1093/oso/9780190618247.001.0001.
- Szczygieł, M. (2017). Konstruktywizm Jeana Piageta i koncepcja zmysłu liczby a edukacja matematyczna. *Edukacja*, 140(1), 7–26. DOI: 10.24131/3724.170101.
- Szczygieł, M., Cipora, K., & Hohol, M. (2015). Liczenie na palcach w ontogenezie i jego znaczenie dla rozwoju kompetencji matematycznych. *Psychologia rozwojowa*, 20(2), 23–33. DOI: 10.4467/20843879PR.15.014.3803.

Marek Lechniak

Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0002-0768-7963>

<https://doi.org/10.35765/slowniki.475>

Logika współczesna i jej rola w naukach ścisłych oraz humanistycznych

Streszczenie

DEFINICJA POJĘCIA: Odróżnia się logikę jako dyscyplinę naukową oraz zastosowania logiki. Logikę można rozumieć w sensie szerszym (semiotyka logiczna, logika formalna i metodologia) i w sensie węższym (logika formalna).

ANALIZA HISTORYCZNA POJĘCIA: Historia logiki sięga starożytności, a jej rozwój następuje nie liniowo, ale w obrębie kolejnych jej postaci. Współczesna logika obejmuje klasyczny rachunek logiczny (wraz z metalogiką i filozofią logiki) i systemy logik nieklasycznych (budowane jako dewiacje logiki klasycznej lub jej rozszerzenia).

UJĘCIE PROBLEMOWE POJĘCIA: Zastosowania logiki w naukach ścisłych obejmują matematykę (w klasycznej matematyce – klasyczny rachunek logiczny) oraz fizykę (w języku matematycznym – klasyczny rachunek logiczny, a w języku wyobrażeniowym – logiki nieklasyczne), w naukach humanistycznych zaś – metodologiczne problemy specyfiki ich przedmiotu (definiowanie ich pojęć i interpretacja humanistyczna jako sposób wyjaśniania oraz stosowanie logik nieklasycznych i ontologii formalnych w charakterystyce pojęć).

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Logika stanowi (przynajmniej *de iure*) podstawę budowy teorii naukowej (precyzja języka teorii, zależność między elementami teorii) oraz przeprowadzanych w niej procedur wyjaśniających i uzasadniających.

Badacze korzystający z osiągnięć współczesnych logik nieklasycznych mogą je wykorzystać w poszukiwaniu metod interdyscyplinarnych w kognitywistyce.

Słowa kluczowe: logika współczesna, logika w naukach ścisłych, logika w naukach humanistycznych, logiki nieklasyczne, kognitywistyka

Definicja pojęcia

Nie można podać jednego określenia tego, czym jest logika. Najogólniej rzecz ujmując, można by przez logikę rozumieć

analizę języka i czynności badawczych w celu podania takich reguł postępowania się językiem i wykonywania owych czynności, które uczyniłyby taką działalność możliwie jak najbardziej skuteczną (Marciszewski, 1970, s. 126).

Należy przy tym odróżnić szersze rozumienie logiki (logika w sensie szerszym, logika ogólna) oraz rozumienie węższe. W ramach pierwszego z nich istnieją dwa kluczowe tematy, wokół których skupia się problematyka logiki: zagadnienie jasnego, konsekwentnego, ścisłego i uporządkowanego myślenia i mówienia oraz zagadnienie poprawnego wnioskowania; z kolei w obrębie rozumienia węższego bada się schematy, czyli formy rozumowań niezawodnych mających najwyższy stopień ogólności, a więc znajdujących zastosowanie w każdej nauce i w języku codziennym (Marciszewski, 1970, s. 133). Rozumienie szersze obejmuje, mówiąc obrazowo, obszar stosowania przysłowka „logicznie” w polskim języku potocznym. Przysłówek ten (podobnie jak gr. *logos*) najczęściej stosuje się do takich czasowników jak „mówi”, „myśli”, „działa”. Logicznie mówi ten, kto w sposób jasny i precyzyjny formułuje swoje myśli, kto umie je przekazać jednoznacznie i zrozumiale. Logicznie myśli ten, kto nie przeczy sam sobie, kto sprawnie umie wyprowadzać wnioski z tego, co już wie, oraz kto myśli poprawnie, to znaczy nie przechodzi od zdań prawdziwych do zdań fałszywych. Logicznie zaś działa ten, kto postępuje w sposób przemyślany i sprawny, kto dobrze stawiając cele działania, osiąga je przy minimum użytych środków. Dlatego zwykle w podręcznikowych ujęciach logiki odróżnia się trzy jej subdyscypliny: logiczną teorię języka, zwaną semiotyką logiczną, logikę formalną oraz metodologię nauk. Przedmiotem semiotyki logicznej jest język, traktowany jako narzędzie sprawnego komunikowania się. Logika formalna dotyczy obiektywnych związków między zdaniami o odpowiedniej formie. Związki te (mające podstawę w relacjach między stanami rzeczy) są wyrażone za pomocą praw logiki, będących gwarantami niezawodności wnioskowań przeprowadzanych w każdej dziedzinie wiedzy. Z kolei ogólna metodologia nauk dotyczy szeroko rozumianych metod postępowania stosowanych w poznawaniu

świata, w szczególności sposobów uzasadniania głoszonych przez nas twierdzeń o świecie (Kamiński, 1992).

Od zarania dziejów myśli odróżnia się dwa spojrzenia na logikę:

- jako dziedzinę wiedzy (średniow. *logica ut doctrina, logica docens*),
- jako narzędzie zdobywania wiedzy, jako zestaw sprawności logicznych (*logica ut disciplina, logica utens*).

Tę dwoistość można już zauważyć w samym ujęciu problematyki logicznej u Arystotelesa; Stagiryta traktował logikę jako propedeutykę wiedzy (nauki), jednak podkreślając, iż chodzi w niej o ogólne prawa naukowe; tymczasem Andronikos z Rodos, porządkując *corpus Aristotelicum*, prace logiczne umieścił na początku zestawu dzieł, nadając im łączną nazwę *organon* (narzędzie). Ta dwoistość rzutuje także na dzisiejsze postrzeganie logiki; z jednej strony mamy zestaw wyrafinowanych systemów naukowych, drobiazgowo badanych za pomocą metod formalnych, z drugiej zaś problem stosowalności tych wyrafinowanych teorii jako narzędzi przydatnych w praktyce tak nauk ścisłych, jak i humanistycznych. Tak logika jako *doctrina*, jak i logika jako *disciplina* odgrywają ważną rolę w interdyscyplinarnej dziedzinie wiedzy, jaką jest kognitywistyka. Logika (i jej subdyscypliny) jest bowiem dyscypliną, która zajmuje się informacją, reprezentacją i przetwarzaniem wiedzy.

Analiza historyczna pojęcia – zarys dziejów logiki

Nie jest łatwo podać jednoznaczną periodyzację dziejów logiki. J.M. Bocheński proponuje wyróżnienie pięciu okresów w historii logiki europejskiej: (a) okres starożytny; (b) okres wczesnośredniowieczny (VII–XI w.); (c) okres scholastyki (XII–XV w.); (d) okres nowożytnej logiki klasycznej (XVI–XIX w.); (e) okres logiki matematycznej. Według Bocheńskiego okresy (b) i (d) były nietwórcze, stąd pozostają do rozważenia jedynie okres starożytny, okres scholastyki i okres logiki matematycznej (Bocheński, 1961). Podkreśla się przy tym, że logika nie rozwijała się w sposób ciągły; nowa logika nie jest kontynuacją logiki starej, gdyż wyrasta z nowych założeń, stosuje inne narzędzia i rozwija problemy wcześniej nieuwzględniane. Rozwój logiki dokonuje się w obrębie postaci (ostrzej postawione problemy, lepsze metody, więcej praw), przy czym postać późniejsza nie musi być doskonalsza od wcześniejszej.

Dotyczy to, zdaniem Bocheńskiego, również logiki matematycznej, która mimo swego technicznego wyrafinowania zatracą subtelność semiotyki średniowiecznej, utożsamiając logikę z rachunkiem logicznym. Według Bocheńskiego można wyróżnić trzy zasadnicze historyczne postaci w uprawianiu logiki, a mianowicie:

- postać antyczną logiki, charakteryzującą się tym, że twierdzenia formułowane są zwykle w języku przedmiotowym, semantyka języka jest słabo rozwinięta, teorie buduje się w uproszczonym języku potocznym ze zmiennymi. Punktem wyjścia badań są wyrażenia formułowane w języku naturalnym, podczas gdy prawa i reguły zostają wyabstrahowane z języka naturalnego;
- postać scholastyczną logiki, w której twierdzenia formułowane są w języku metapredmiotowym, semantyka jest wnikliwa i rozbudowana. Teoria budowana jest w języku potocznym, bez zawężania funkcji semiotycznych, a prawa i reguły są rezultatem abstrakcji z języka naturalnego;
- postać matematyczną logiki charakteryzującą się nawrotem do starożytności; twierdzenia w niej są wyrażane w języku przedmiotowym z szerokim zastosowaniem zmiennych przy ścisłym określaniu semiotycznym znaków. Semantyka języka (przed latami trzydziestymi XX wieku) była uwzględniana w niewielkim zakresie. Nowością jest zastosowanie języka sztucznego, przy czym dominuje w tej postaci podejście konstruktywne do logiki polegające na pierwszeństwie formalizmu języka przed jego interpretacją. Na takim języku opierają się systemy sztucznej inteligencji wykorzystywane w kognitywistyce do modelowania procesów poznawczych w umyśle.

Ujęcie problemowe pojęcia

Subdyscypliny logiki

Podziału logiki na subdyscypliny nie da się przeprowadzić w sposób ostry. W typowym jednak podejściu wyróżnia się trzy jej działy, a mianowicie:

- semiotykę logiczną, w której obrębie odróżnia się dociekania na płaszczyźnie syntaktycznej (relacje: wyrażenie – wyrażenie),

semantycznej (gdzie bada się relacje typu: wyrażenie języka – obiekt rzeczywistości, do którego odnosi się wyrażenie) oraz pragmatycznej (dociekania nad relacjami pomiędzy wyrażeniami języka a ich użytkownikami). Badania semiotyczne ogólnej teorii języka i jako takie stanowią element analiz kognitywistycznych, umożliwiając badanie powstawania znaczenia w percepcji, myśli i języku jako systemie znaków;

- logikę formalną, badającą związki formalne między wyrażeniami języka w aspekcie formalnej poprawności; chodzi o to, jak z wyrażzeń o określonej strukturze wynikają wyrażenia o takiej, a nie innej strukturze;
- metodologię nauk, badającą szeroko rozumiane metody naukowe, prowadzące do wartościowych poznawczo rezultatów.

Zakresem typowo określonej semiotyki logicznej jest następujący zespół zagadnień: teorie kategorii składniowych (lub inne rodzaje gramatyk logicznych), zagadnienia związane z różnymi rodzajami podziałów logicznych (w tym różne teorie klasyfikacji i typologii); analiza błędów w słownym wyrażaniu myśli, dociekania nad definicjami, problemy semantyczne, takie jak zagadnienia denotowania (np. denotowania przez nazwy jednostkowe i deskrypcje) czy problematyka analizy pojęcia prawdziwości w językach sformalizowanych i naturalnych; pragmatyczne koncepcje i problemy używania języka (np. problematyka znaczenia, uznawania, problemy implikatury konwersacyjnej etc.).

Z kolei w obrębie logiki formalnej konstruowane są różnego rodzaju systemy (Marciszewski, 1987), badane są formalne własności tych systemów oraz relacje między systemami formalnymi (metalogika), a także prowadzone są dociekania filozoficzne nad naturą i stosowalnością owych systemów (filozofia logiki).

Problematyka metodologii nauk (często współcześnie zamiast tej nazwy używa się terminu „filozofia nauki”) obejmuje ogólne dociekania nad naturą nauki traktowanej w sposób czynnościowy oraz nad nauką jako wytworem, czyli teorią naukową; do problematyki metodologii zaliczyć można także teorię pytań, rozmaite teorie rozumowań oraz argumentacji, dociekania nad strukturą i dynamiką teorii naukowych.

Współczesna logika formalna

Nazwa „logika formalna” bierze się stąd, że przedmiotem tego działu logiki jest formalna strona wnioskowań. To znaczy, że w logice formalnej, mówiąc z grubsza, abstrahuje się od treści wyrażen, koncentrując się na ich kształtach. Dlatego najczęściej w dziale tym mówi się o formach zdaniowych, a nie o zdaniach, jako że nie chodzi w nim o wartość logiczną poszczególnych zdań, ale o związki między wyrażeniami określonej postaci a innymi wyrażeniami określonej postaci. Charakterystyczną (choć niekonieczną) cechą współczesnej logiki formalnej jest też jej symboliczność: jej prawa (reguły) są wyrażone w języku symbolicznym, gwarantującym jednoznaczność sformułowań i mechanizowalność procedur uzasadniania. Język symboliczny oraz rachunkowy („matematyczny”) charakter sprawiają, że systemy logiki formalnej, które stanowią bazę dla uzasadnienia jej twierdzeń, są wzorami ścisłości i precyzji naukowej (Łukasiewicz, 1961; Borkowski, 1970).

Klasyczny rachunek logiczny

Podstawowy charakter w logice formalnej ma klasyczny rachunek logiczny obejmujący klasyczny rachunek zdań oraz węższy rachunek predykatów (klasyczny rachunek kwantyfikatorów) z identycznością. Niektórzy do logiki klasycznej włączają dodatkowo algebrę zbiorów i relacji. Przy odpowiednich założeniach teoria zdań kategoriycznych może być traktowana jako część węższego rachunku predykatów; podobnie do logiki klasycznej zwykło się zaliczać tzw. systemy Stanisława Leśniewskiego (zwłaszcza prototetykę i ontologię).

Klasyczny rachunek zdań może być zbudowany na różne sposoby, wśród których (na poziomie syntaktycznym) można wskazać system aksjomatyczny lub system dedukcji naturalnej (założeniowy lub system Gentzena); w wersji semantycznej może to być np. użycie tabel 0-1 lub tablic analitycznych. Niezależnie od sposobu budowy można wskazać, że klasyczny rachunek zdań opiera się na dwóch podstawowych założeniach, a mianowicie zasadzie dwuwartościowości (każde zdanie jest albo prawdziwe, albo fałszywe) oraz założeniu, że każdy funkcyj jest prawdziwościowy (tzn. że istnieje jednoznaczna relacja między wartością

logiczną jego argumentów a wartością logiczną całego wyrażenia przezeń utworzonego). Każde z tych założeń zakłada pewną idealizację języka naturalnego, w którym nie każde wyrażenie jest semantycznie zdeterminowane do prawdziwości czy fałszywości (np. wypowiedzi niepełne), jak i istnieje wiele spójników, które nie mają charakteru funktorów prawdziwościowych (funktory nieekstensjonalne).

System dedukcji naturalnej może przyjąć postać np. ujęcia założeniowego klasycznego rachunku zdań, które opiera się na dwóch typach reguł – regułach tworzenia założeniowych dowodów nie wprost oraz pierwotnych regułach dołączania nowych wierszy do dowodu określających formalne zachowanie podstawowych spójników logicznych (negacja, koniunkcja, alternatywa, implikacja czy równoważność). System ten może być rozbudowany do węższego rachunku predykatów z identycznością przez odpowiednie wzbogacenie słownika, dodanie reguł składowania dla nowych stałych oraz dodanie reguł określających formalne zachowanie kwantyfikatorów i znaku identyczności.

Logiki nieklasyczne

Wskazuje się, że klasyczny rachunek logiczny może być niewystarczający dla adekwatnego wyrażenia różnego rodzaju zwrotów występujących w języku naturalnym i opartym na nim języku nauk szczegółowych. Na poziomie wypowiedzi, do których analizy wystarczy rachunek zdań, taką racją do modyfikacji klasycznego rachunku zdań (KRZ) może być z jednej strony fakt występowania w języku naturalnym licznych zwrotów o charakterze nieekstensjonalnym (jak np. zwrotów modalnych), z drugiej zaś fakt, że kwestionuje się ogólną ważność niektórych praw klasycznego rachunku zdań; w pierwszym wypadku prowadzi to do rozszerzeń języka klasycznego rachunku zdań o zwroty nieekstensjonalne, podczas gdy w drugim – do tzw. logik dewiacyjnych, gdzie zbiór symboli nowego systemu jest taki sam jak zbiór symboli KRZ, podczas gdy pewne odpowiedniki też KRZ nie są tezami owych nowych logik.

Logiki modalne

Wśród rozszerzeń KRZ można wskazać systemy logik zdań modalnych aletrycznych (zwroty: „jest konieczne, że”, „jest możliwe, że”), logik epistemicznych (zwroty: „wie, że”, „jest przekonany, że”, „uznaje, że”), logik deontycznych („jest nakazane, że”, „jest dozwolone, że”, „jest zakazane, że”), logik temporalnych, logik zdań kauzalnych etc. Wskazuje się (von Wright, 1951) liczne analogie formalne co do zachowań funktorów tych logik (np. funktory „jest konieczne, że”, „wie, że”, „jest nakazane, że”), ale są też pewne dysanalogie (np. chociaż zdanie „jest konieczne, że p ” w wielu systemach implikuje p , to jednak tak nie jest dla zwrotu „jest nakazane, że p ”). Ze względu na owe analogie często nazywa się logikami modalnymi (w sensie szerszym) wszystkie rozszerzenia logiki klasycznej. Istnieje kontinuum systemów logik nieklasycznych. Warto wskazać, że już w połowie XX wieku wskazywano na trudności intuicyjne w rozumieniu syntaktycznie sformułowanych pewnych praw logik modalnych (np. praw iterowania zwrotów modalnych (von Wright, 1951; Świrydowicz, 2004; Hughes, & Cresswell, 1968)). Jednak fakt znalezienia (pod koniec lat pięćdziesiątych XX wieku) prostych formalnych semantyk dla logik modalnych (np. semantyki relacyjne, zwane semantykami Kripkego) doprowadził do nowej fazy w rozwoju tych logik. Semantyki relacyjne wychodzą z idei Leibniza, że zdanie jest konieczne, jeśli jest prawdziwe w każdym możliwym świecie. *Novum* pozwalające zastosować tę ideę do różnych systemów logik modalnych polega na tym, że ową kwantyfikację światów możliwych ogranicza się przez wprowadzenie relacji dostępności między światami; w takim ujęciu zdanie jest konieczne, gdy jest prawdziwe w każdym świecie możliwym, dostępnym ze świata aktualnego. Dzięki idei światów możliwych okazało się, że liczne własności funktorów modalnych prowadzące do konstruowania różnych systemów logik modalnych na poziomie semantyki formalnej znajdują swój wyraz w różnych własnościach relacji alternatywności (dostępności) między światami możliwymi – takie semantyki stanowią narzędzie rozstrzygnięcia, czy pewne formuły logik modalnych są tezami określonych systemów logik modalnych, czy nie.

Logiki wielowartościowe

Z drugiej strony wskazuje się liczne systemy, które można by traktować jako dewiacje logiki klasycznej. Najbardziej znane z nich to systemy logik wielowartościowych; przyjmuje się w nich, że każdy funktor powinien być traktowany jako funkcja; zachowując charakterystykę matrycową funktorów, zwiększa się liczbę wartości tych matryc, rozszerzając liczbę możliwych charakterystyk funktorów (np. już w logice 3-wartościowej możliwych jest 27 funktorów jednoargumentowych i 3^9 funktorów dwuargumentowych). W pierwszej fazie rozwoju tych logik zdawało się, że ze względu na rewizję zasady dwuwartościowości są one konkurencyjne względem KRZ (nie obowiązywały w nich np. prawa niesprzeczności czy wyłączonego środka). Jednak w następnej fazie rozwoju tych logik odkryto, że można zachować matrycową charakterystykę funktorów bez rewizji klasycznego pojęcia prawdziwości, a matryce wielowartościowe mogą być tworzone przez krzyżowanie podziału zdań na prawdziwe – fałszywe z innymi podziałami zdań (np. zdeterminowane – niezeterminowane, poinformowany – niepoinformowany etc.); przykładem mogą tu być matryce dla systemów logik relewantnych, w których to logikach chodzi o zachowanie jakiegoś związku treściowego (tzw. kryterium relewancji) między przesłankami a wnioskiem (takiego kryterium nie spełnia logika klasyczna choćby ze względu na obowiązujące w niej prawo przepełnienia: $p \wedge \neg p \rightarrow q$). Niewątpliwą korzyścią ze stosowania charakterystyki matrycowej jest fakt, że dostarcza ona prostego kryterium rozstrzygalności, jak również prostego sposobu charakterystyki nowych funktorów; niestety minusem bywa czasami nieintuicyjność otrzymanych wyników.

Zastosowania logiki w naukach ścisłych

Interdyscyplinarny charakter logiki polega nie tylko na tym, że łączy ona różne aspekty przedmiotów badanych w ramach poszczególnych dyscyplin (w naszym przypadku w naukach ścisłych i humanistycznych). Ponadto umożliwia poszukiwanie elementów wspólnych dla różnych nauk, które z punktu widzenia danych dyscyplin różnią się od siebie.

Zastosowania logiki w matematyce

Wskazuje się, że w matematyce występują trzy zasadnicze typy pojęć (Grzegorzcyk, 1981):

1. Pojęcia logiczne, czyli przede wszystkim funktory prawdziwościowe, kwantyfikatory (zwroty określające, dla jakiej ilości przedmiotów zachodzi zdanie w tych zwrotach wypowiedziane; dla każdego x , dla pewnego x), a także pojęcie identyczności wyrażone w zwrocie $x=y$.
2. Pojęcia teoriomnogościowe, takie jak zbiór, przynależność elementu do zbioru, ogólne pojęcie relacji, ogólne pojęcie funkcji.
3. Pojęcia arytmetyczne: zbiór liczb naturalnych, zbiór liczb rzeczywistych, działania arytmetyczne: dodawanie, mnożenie i inne dające się określić za pomocą powyższych.

Rola pojęć każdej z trzech wymienionych grup w klasycznej matematyce może być najkrócej scharakteryzowana w ten sposób, że logika dostarcza zasadniczo metod konstrukcji pojęć i budowy dowodów, teoria mnogości dostarcza pojęć najogólniejszych (np. pojęcia zbioru), zaś pojęcia arytmetyczne są podstawą dla treści specyficznie matematycznych.

Logika wychwytuje sposoby wnioskowania stosowane w naukach i uznawane za poprawne oraz tworzy z nich systemy logiczne. Systemy te są zbiorem praw i reguł, do których się stosując, można odtworzyć te wszystkie wnioskowania, które spontanicznie uznajemy za bez zarzutu. Pod względem uznawania lub nieuznawania pewnych sposobów wnioskowania istnieje w matematyce duża zgodność poglądów. Nie ma takiej sytuacji, ażeby pewne reguły wnioskowania jedni uważali za poprawne, a inni za całkowicie błędne. Są tylko różnice w rozumieniu pewnych reguł i w poglądzie na zakres ich stosowania. Zbiór, praktycznie biorąc, wszystkich metod wnioskowania spotykanych w matematyce, daje tzw. klasyczny system logiki (Grzegorzcyk, 1981, s. 65).

Idąc za Grzegorzcykiem, którego głos wyraża raczej powszechne poglądy w tej kwestii, można stwierdzić, że we wnioskowaniach przeprowadzanych w klasycznej matematyce wystarcza klasyczny rachunek logiczny i nie ma potrzeby wprowadzać innych rodzajów logik (np. nieklasycznych).

Pewna komplikacja w tej sprawie nastąpiła jednak na początku XX wieku wraz z programem intuicjonizmu. Według zwolenników tego

poglądu (L.E.J. Brouwer, A. Heyting) klasyczna matematyka nie powinna opierać się na logice klasycznej, której dowody charakteryzuje niekonstruktywność, będąca pochodną przyjęcia w logice klasycznej prawa wyłączonego środka $p \vee \neg p$ i prawa podwójnego przeczenia $p \equiv \neg \neg p$; na podstawie pierwszego z nich zachodzi stan rzeczy stwierdzony przez zdanie albo stan rzeczy stwierdzony przez negację tego zdania, podczas gdy na podstawie drugiego zdanie jest równoważne negacji negacji tego zdania. Tymczasem intuicjonizm, wychodząc z innej niż logika klasyczna perspektywy badawczej, inaczej rozumie wyrażenie zdaniowe oraz negację; zmienna zdaniowa reprezentuje bowiem stan rzeczy (opis obiektu matematycznego) wraz z procedurą jego skonstruowania; wtedy oczywiście prawo wyłączonego środka interpretowane jako „każdy stan obiektu matematycznego jest skonstruowany lub negacja tego stanu jest skonstruowana” nie może być przyjęte; w każdym bogatszym rachunku istnieją bowiem zdania takie, że nie są one udowodnione ani ich negacja nie może być udowodniona; podobnie z prawem podwójnej negacji – w odczytaniu intuicjonistycznym brzmi ono: „absurdalność absurdalności zdania implikuje jego prawdziwość” (konstruowalność), co dla intuicjonistów jest nie do przyjęcia. Zatem w programie intuicjonistycznym logika staje się częścią matematyki, a mianowicie wyabstrahowaną z faktycznej działalności matematyków teorią procedur dowodowych stosowanych w matematyce. Rachunkiem służącym kontrolowaniu rozumowań w matematyce intuicjonistycznej jest logika intuicjonistyczna, która zalicza się do logik nieklasycznych (przy interpretacji intuicjonistycznej – konkurencyjnych względem logiki klasycznej; przy interpretacji bardziej precyzyjnej logika ta nie jest konkurencyjna, ale po prostu dotyczy innej, nieontologicznej, perspektywy badawczej).

Stosowanie logiki w fizyce

Wskazuje się (np. idąc za W. Heisenbergiem), że przez długi czas uważano, iż problem języka w naukach przyrodniczych nie odgrywa większej roli. Jednak, jak stwierdza Heisenberg, sytuacja uległa zasadniczej zmianie na początku XX wieku w związku z odkryciem teorii względności i mechaniki kwantowej (zwłaszcza odkryciem ciał promieniotwórczych);

ta przemiana spowodowała radykalną zmianę założeń dotyczących budowy materii i procesów jej przemiany.

Fizycy XX w. musieli przyzwyczać się m.in. do tego, aby nie rozpatrywać mikrokosmosu cząstek elementarnych na obraz i podobieństwo makrokosmosu dużych ciał, które są bezpośrednio dostępne zmysłom. Cząstkom elementarnym, przynajmniej do pewnego stopnia, nie mogą już być przypisywane nawet własności geometryczne i kinematyczne, jak kształt i ruch w przestrzeni. Cząstka elementarna fizyki współczesnej jest jeszcze daleko bardziej abstrakcyjna niż atom Demokryta, choć w ujęciu Greków atomy już nie miały takich własności jak zapach, kolor, smak, lepkość, temperatura. Fizyk niemiecki dodaje jeszcze tę myśl, że chcąc wyrobić sobie obraz cząstek elementarnych, nie możemy już w zasadzie abstrahować od procesów fizycznych, dzięki którym dowiadujemy się o tych cząstkach (Kiczuk, 2001, s. 143).

W związku z wyżej nakreśloną złożonością pojęciową fizyki współczesnej Heisenberg zauważa, że fizycy posługują się dwoma rodzajami języka. Z jednej strony prawa fizyki stwierdzające określone zależności zachodzące w przyrodzie są wyrażone w języku matematycznym, w postaci równań tego języka. Owe równania teorii fizykalnych podają matematyczny obraz różnego typu zdarzeń. Z drugiej strony terminy teorii fizykalnej określone przez równania tych teorii (np. w mechanice Newtona: czas, masa, przestrzeń, prędkość, siła, położenie, przyspieszenie, ciało) mają pogładową treść zjawiskową i mogą wchodzić do języka potocznego. Kolejne teorie fizykalne wprowadzają nowe pojęcia za pomocą praw wyrażonych w języku matematycznym, które trafiają znowu do języka potocznego (np. termodynamika – pojęcie entropii), wprowadzając do niego nowe treści wyobrazeniowe. Dlatego, zdaniem Heisenberga, nie można unikać mówienia o drugim (poza matematycznym) języku teorii fizykalnych, a mianowicie języku wyobrazeniowym; ten ostatni jest odpowiednio uściślonym językiem potocznym.

Fizyka, zdaniem Heisenberga, składa się nie tylko z eksperymentowania i mierzenia po jednej stronie i z aparatu matematycznego – po drugiej stronie. Do obowiązków fizyka należy również wyjaśnienie za pomocą języka potocznego wszystkiego, co właściwie dzieje się w tej grze między eksperymentem i matematyką (Kiczuk, 2001, s. 148).

W związku z użyciem języków tych dwóch rodzajów pozostaje konieczność użycia dwóch języków logiki do kontrolowania wnioskowań

fizykalnych. Po stronie aparatu matematycznego jest to język klasycznego rachunku logicznego, a po stronie języka wyobraźniowego – użycie logik nieklasycznych dla determinacji pojęć i rozumowań przeprowadzanych w tym języku.

Wskazywano, że różne systemy logik nieklasycznych mogą być przydatne w zastosowaniu do wyrażalności też wyobraźniowego języka fizyki. Mogą to być mianowicie logiki zdań czasowych (w szczególności logiki tensalne, czyli logiki czasów gramatycznych, wyrażające w sposób ścisły różne własności czasu oraz logiki interwałów czasowych czy też chronologie), logiki związku przyczynowego, jak również różnego rodzaju logiki kontrfaktycznych okresów warunkowych (ze względu na fakt, że prawa fizyki zwykle wyrażane są za pomocą tego typu zdań). Próbowano także konstruować szczególne systemy logik nieklasycznych, których funkctory miały być adekwatne do wyrażalności osobliwych stanów rzeczy stwierdzanych w fizyce (np. logiki mechaniki kwantowej) czy też stosować do nich istniejące systemy (np. logik wielowartościowych).

Stosowanie logiki w naukach humanistycznych

Precyzyjne omówienie stosowania logiki w naukach humanistycznych wymagałoby najpierw zdeterminowania tego, co rozumiemy przez nauki humanistyczne. Najogólniej przyjmuje się, że stanowią one zespół dyscyplin, zajmujących się różnymi przejawami i wytworami życia psychiczno-duchowego człowieka (Kamiński, 1992); tak określone rozpadają się na co najmniej trzy grupy: o człowieku jednostkowym i społeczeństwie (np. psychologia, antropologia kulturowa nauki społeczno-ekonomiczne), nauki o wytworach kulturowych (np. nauki prawnicze, o religii i moralności czy o sztuce i o języku), nauki o dziejach człowieka żyjącego w społeczeństwie oraz jego wytworów (nauki historyczne). Dodatkowo pojawia się dzisiaj pytanie, czy filozofię traktować jako naukę humanistyczną (jak sugeruje choćby przyjęty w Polsce podział dyscyplin naukowych) czy też nie. Trudno bowiem oddzielić dyskurs filozoficzny od zakładanej logiki; inaczej natomiast może być w nie-filozoficznych naukach humanistycznych.

Logika w filozofii

Rola logiki w dyskursie filozoficznym jest nie do przecenienia. Często dyskurs filozoficzny stanowi niejako wzorzec dyskursu naukowego. Występuje jednak wielka różnorodność sposobów uprawiania filozofii, dopuszczalnych źródeł poznania i metod wyjaśniania i uzasadniania filozoficznego. Rola logiki w wyjaśnianiu i uzasadnianiu twierdzeń zależy od przyjętej opcji. W pewnych typach filozofii (zwłaszcza w filozofii systematycznej) dominuje uzasadnienie dedukcyjne twierdzeń, a to opiera się na niezawodnych schematach wnioskowania. Twierdzenia filozofii stanowią często logiczne (czy treściowe) konsekwencje (uzasadnionych w sposób bezpośredni, np. na podstawie intuicji intelektualnej) zdań bazowych; są więc twierdzenia filozoficzne uzasadnione w sposób pośredni, tzn. na podstawie argumentacji, w której schematy niezawodnego wnioskowania odgrywają zasadniczą rolę (Janeczek, 2018). Przy tym dyskurs filozoficzny prowadzony jest w języku naturalnym (czasem odpowiednio technicznym). Zagadnienia związane z argumentacją są również istotne dla kognitywistów, gdyż dotyczą człowieka jako bytu cielesno-umysłowego. Dodatkowo tworzenie algorytmów zdolnych do argumentacji należy do obszaru sztucznej inteligencji. Logika formalna może także znaleźć (i znajduje) zastosowanie w formalnych rekonstrukcjach wnioskowań filozoficznych, mających za cel znalezienie założeń, *implicite* zawartych w podstawach systemu filozoficznego oraz badaniu formalnej poprawności tak zrekonstruowanych wnioskowań. Nierzadko taka analiza prowadzi do wskazania potrzeby zastosowania jakichś systemów logik nieklasycznych lub do wykazania, że aby uzasadnić pewne tezy, trzeba środków logicznych o takiej, a nie innej sile inferencyjnej.

Z kolei niektórzy (zwłaszcza XX-wieczni) filozofowie postulują ograniczenie aspiracji poznawczych filozofii, podkreślając, że winna się ona skoncentrować na dobrze uzasadnionej analizie pojęciowej (filozofia analityczna, Bocheński, 1993). Filozof analityczny, korzystając z narzędzi logiki, przeprowadza analizy językowe (przy czym konsekwencją logiczności jest prostota i jasność) oparte na badaniu sposobów używania słów języka naturalnego, aby ująć ich znaczenie; takie podejście zakłada paralelizm język – rzeczywistość (ta ostatnia znajduje odzwierciedlenie w języku). Podstawą badania pojęcia ma być uchwycenie jego cech konstytutywnych i konsekwentnych oraz określenie relacji

tego pojęcia (jako bytu abstrakcyjnego) do innych pojęć. Do uchwycenia cech pojęcia ma służyć analiza użyć odpowiadającego mu wyrazu języka naturalnego, uzupełniona o element regulujący, czyli ów moment konstrukcyjny (Łukasiewicz, 1961). Ze względu na logiczny charakter analiz filozofia przyjmuje postać logiki stosowanej. Logika traktowana jest w takich analizach jako najogólniejsza ontologia, „teoria najogólniejszych, najprostszych związków między wszelkimi rzeczami w ogóle, a dzięki temu doskonałym narzędziem nie tyle rozumowania, ile właśnie analizy” (Bocheński, 1993, s. 191). Tak rozumiana oparta na logice analiza logiczna zmierza do definicji analizowanego pojęcia i uchwyceniu jego podgatunków oraz relacji między nimi; może ona stać się z kolei narzędziem do wykorzystania w uprawianiu nauk humanistycznych.

Na taką analizę składają się zwykle następujące czynności: odwołanie się do znaczenia analizowanego wyrażenia w języku naturalnym; porządkowanie intuicji związanych z użyciem stosownego pojęcia języka naturalnego polegającego na odróżnianiu różnych znaczeń wyrazu przez odróżnianie różnych kontekstów jego użycia; charakterystyka rodzaju najbliższego definiowanego gatunku; określenia formalnych własności pojęcia (własności, relacji); wprowadzanie podziału analizowanego pojęcia poprzedzone poszukiwaniem zasady podziału; analiza poszczególnych członów podziału; zaczyna się od prób określenia istoty każdego z członów podziału oraz podania własności tak określonych pojęć (Bocheński, 1993). Stosowane są w analizie tradycyjne środki logiczne: kategoryalizacja pojęć (dziedziny rzeczywistości, której dotyczy analiza), definiowanie oraz podział logiczny. Działania te wsparte są często symbolizacją (a nie formalizacją) oraz stosowaniem narzędzi formalnych logiki I rzędu. Wiele efektów zastosowania tych dość prostych środków przez Bocheńskiego zostało zauważonych w literaturze (np. prace dotyczące wolności, autorytetu, przedsiębiorstwa przemysłowego etc.).

Stosowanie logiki w szczegółowych naukach humanistycznych

Często, poszukując cech świadczących o odrębności nauk humanistycznych od nauk przyrodniczych, wskazuje się ich odrębny przedmiot o charakterze kulturowym (charakteryzujący się dwiema cechami:

kultura składa się z działań i wytworów psychiki ludzkiej – ujmowanych tak w perspektywie indywidualnej, jak i społecznej – podczas gdy rzeczywistość ujmowana przez nauki przyrodnicze istnieje niezależnie od takich operacji oraz kultura cechuje się momentem wartości, związanym z działaniem i wytworami ludzkimi). Ten kulturowy charakter rzutuje na ujęcie przedmiotu nauk humanistycznych jako indywidualnego – zjawiska kulturowe ze względu na ich powiązanie z działalnością konkretnych osób i grup społecznych noszą znamiona jednostkowości, niepowtarzalności i czasowości.

Owa indywidualność przedmiotu pociąga szereg trudności logicznych w jego charakterystyce. Mogą to być mianowicie:

- logiczna osobliwość pojęć humanistyki (nieostrość i niewyraźność, wieloznaczność) i trudności w ich normalnym definiowaniu (pojęcia typologiczne, pojęcia mające rodziny znaczeń, charakter perswazyjny pojęć); potrzeba stosowania definicji cząstkowych, definicji perswazyjnych (Pawłowski, 1986);
- odmienny typ wyjaśniania (interpretacja humanistyczna).

Logiczne problemy definiowania w naukach humanistycznych

Ze względu na specyfikę pojęć humanistyki szczególną rolę w eksplikacji tych pojęć odgrywa logiczna teoria definicji. Szczególną cechą tych pojęć jest fakt, że w przeważającej mierze mają one charakter typologiczny albo rodziny znaczeń (Pawłowski, 1986).

W pierwszym wypadku granice ich zakresów są nieostre (albo treść tych pojęć jest niewyraźna); można by powiedzieć, że pojęcia typologiczne, odnoszące się często do cech stopniowalnych (inteligentny, pracowity), mają charakter porządkujący, a nie klasyfikujący. W humanistyce najczęściej wyznacza się pojęcia typologiczne na podstawie empirycznie danych uporządkowań. W typologii możemy wykorzystać np. uporządkowanie według typów krańcowych (idealnych), wskazujących na minimalny i maksymalny punkt skali własności stopniowalnych.

Z kolei pojęcia mające rodziny znaczeń (np. sztuka, awangarda, powieść, tragedia) charakteryzują się tym, że ich zakresem nie jest zbiór przedmiotów posiadających zespół cech wspólnych wszystkim elementom zakresu, ale szereg podzbiorów przedmiotów połączonych

częściowymi podobieństwami. Tej rodzinie podzbiorów odpowiada rodzina znaczeń, stanowiąca znaczenie danego pojęcia. Wskazywano (Wittgenstein), że takie pojęcia nie są możliwe do zdefiniowania; jednak należy zmodyfikować to stanowisko – po prostu nie można dla takich pojęć dać definicji równościowej (*per genus proximum et differentiam specificam*, natomiast środkiem do ich definiowania jest zespół definicji częściowych, podających częściowe kryteria stosowalności pojęcia (przez wskazanie warunku koniecznego lub też wystarczającego); charakterystyczną cechą pojęć definiowalnych jedynie częściowo jest to, iż żadna liczba definicji częściowych nie daje w sumie definicji pełnej – zakres takiego pojęcia pozostaje otwarty (czyli zawsze będą istniały przedmioty, co do których pytanie, czy należą do zakresu tego pojęcia, pozostanie otwarte).

Interpretacja humanistyczna

Indywidualność przedmiotu nauk humanistycznych domaga się przyjęcia innego sposobu wyjaśniania niż eksplanacja przeprowadzana w naukach przyrodniczych. Ta ostatnia, zgodnie z określeniem podanym przez Ajdukiewicza, stanowi proces myślowy, w którym dla zdania przyjętego szukamy racji tylko wśród zdań przyjętych (wyjaśnianie uniwersalizujące) lub wśród zdań przyjętych i nieprzyjętych (wyjaśnianie teoretyczne). W rozumowaniu zwanym wyjaśnianiem kierunku wnioskowania jest odwrotny w stosunku do kierunku wynikania logicznego (bo wnioskiem jest poszukiwana racja stwierdzanego stanu rzeczy), zdanie otrzymane w punkcie dojścia wyjaśniania nie jest zdaniem pewnym, a jedynie prawdopodobnym; wyjaśnianie bowiem przebiega według schematu wnioskowania redukcyjnego (od następstw do racji). Natomiast w wyjaśnianiu stosowanym w naukach humanistycznych w charakterze naczelnych przesłanek występują zdania bezpośrednio oparte na rozumieniu czyichś wypowiedzi lub czyjegoś zachowania albo też istnienia wytworów czyjejs działalności. Logiczna analiza sposobu wyjaśniania w naukach humanistycznych doprowadziła do koncepcji interpretacji humanistycznej. Jest to (najogólniej mówiąc) czynność formułowania odpowiedzi na pytanie typu: „Dlaczego x podjął taką a taką czynność?”, albo: „Dlaczego x nadał swojemu wytworowi

taką a taką postać?”. Odpowiedź ta składa się z: 1) założenia o racjonalności x -a, 2) opisu wiedzy x -a wyodrębniającego możliwe do podjęcia przez niego czynności oraz określającego ich rezultaty, 3) opisu charakterystycznego dla x -a porządku wartości utworzonego z owych rezultatów, takiego, że rezultat rozważanej czynności jest wartością preferowaną. Podstawowe dla tego typu wyjaśniania jest założenie o racjonalności podmiotu działającego (dla czynności podejmowanej w warunkach pewności) stwierdzające, że

Jeśli x ma podjąć jedną z czynności c_1, \dots, c_n , które na gruncie jego wiedzy wykluczają się i dopełniają oraz niezawodnie prowadzą do rezultatów s_1, \dots, s_m , a przy tym rezultaty te są uporządkowane przez charakterystyczną dla x -a relację preferencji, to x podejmie czynność c_i związaną z najwyższym preferowanym rezultatem s_j (Kmita, 1971).

Oczywiście, jak podkreślają autorzy, interpretacja humanistyczna nie jest jedynym sposobem wyjaśniania stosowanym w naukach humanistycznych (inne to np. wyjaśnianie genetyczne), niemniej pozwala w jakimś stopniu ująć specyfikę przedmiotu nauk humanistycznych.

Logiki nieklasyczne i ontologie formalne

Jak wskazywano, przedmiot nauk humanistycznych ze względu na wysoki stopień skomplikowania jest trudny do precyzyjnego ujęcia – większość pojęć „wymyka się” definicji normalnej, zmuszając badaczy do innego rodzaju zabiegów poznawczych. Na tym tle pojawia się możliwość stosowania różnego rodzaju logik nieklasycznych, stanowiących niejako definicje przez postulaty różnych pojęć związanych z poznaniem ludzkim (np. logiki epistemiczne), jego działaniem (np. logiki norm, logiki deontyczne) czy wyborami (np. logiki preferencji).

Jako przykład mogą tu służyć logiki wiedzy i przekonań oraz formalne teorie zmiany przekonań czy dynamiczne logiki przekonań (Lechniak, 2011). Są to najczęściej systemy normalnych logik modalnych, z funktorami modalnymi interpretowanymi jako zwroty „ x jest przekonany, że”, „ x akceptuje, że”, „ x wie, że”, „w obliczu wiedzy o stanie rzeczy p , x zmienia swoje przekonanie” etc. Na przykład kolejne aksjomaty logiki modalnej przy interpretacji epistemicznej mogą być interpretowane jako tezy

stwierdzające własności wiedzy (prawdziwość wiedzy, niesprzeczność wiedzy, nieomyślność w zakresie własnej wiedzy, zupełność w zakresie własnej wiedzy, itp.). Interesujące rezultaty daje również kombinowanie różnych systemów logik nieklasycznych, np. logiki wiedzy z logiką przekonań, logiki przekonań z logiką dynamiczną (zmiany); choć każdy z systemów opiera się na prostych aksjomatach, to jednak kombinacja pojęć opisywanych przez aksjomaty jednego z systemów z tymi opisywanymi przez aksjomaty innego systemu ukazuje złożoność pojęć dotyczących ludzkiej wiedzy i przekonań (np. logiki oznajmień publicznych, w których stan przekonaniowy podmiotu (grupy podmiotów) jest pochodną uzyskania informacji ze spostrzeżeń podmiotu i informacji uzyskanych od innych podmiotów).

Podobnie do logik nieklasycznych ciekawy wkład do analizy pojęć nauk humanistycznych, a w szczególności pojęć kognitywistyki, mogą pełnić ontologie formalne. Są to nabudowane najczęściej na logice klasycznej – ewentualnie wzbogaconej np. przez mereologię (teorię zbiorów w sensie kolektywnym) – systemy formalne (aksjomatów i definicji), których celem jest formalna charakterystyka pojęć jakiejś dziedziny przedmiotowej. Modele uzyskiwane tą drogą mogą służyć jako wyrażona w języku formalnym konceptualizacja dziedziny wiedzy będącą podstawą budowania bazy wiedzy. Ontologie różni od systemów logiki formalnej fakt, że ich celem nie jest określenie relacji konsekwencji dla systemu zdań (jak w logice), ale raczej dostarczenie możliwie adekwatnego formalnego modelu dziedziny. Stąd takie ontologie mogą składać się z kilkudziesięciu aksjomatów i definicji (inaczej niż w systemach logiki, które zwykle składają się z niewielkiej liczby aksjomatów). Jako przykład można podać ontologię przekonań, ontologie obiektów historycznych etc. Proponowano też (np. Jerzy Perzanowski) zastosowanie ontologii formalnych w kognitywistyce (*psychoontologie*).

Refleksja systematyczna z wnioskami i rekomendacjami

Odpowiednio specjalistyczny język oraz operacje wiedzotwórcze, takie jak wyjaśnianie uzasadnianie twierdzeń, odgrywają zasadniczą rolę w każdej nauce. Stąd logika, jako teoria zarówno języka jako narzędzia

sprawnego formułowania i komunikowania myśli, jak i formalna teoria rozumowań, winna odgrywać podstawową rolę jako dostarczająca wewnętrznych zasad organizowania teorii naukowych (rachunek logiczny jako podstawa rozumowań i argumentacji naukowych) oraz baza zewnętrzna (logika jako narzędzie do prowadzenia metarefleksji nad teorią naukową) w każdej dyscyplinie naukowej. Dotyczy to tak nauk ścisłych, jak i humanistycznych (choć za każdym razem jej rola może odnosić się do innych problemów). W ramach kognitywistyki poszukuje się metodologii łączącej te dwie dziedziny nauk. Zadaniem logiki/logik jest precyzowanie zasad odróżniania rozumowań poprawnych od niepoprawnych, co kognitywistyka lokuje w obszarze reprezentacji rozumowań. Przetwarzanie informacji w ludzkim umyśle i w leżącym u jego podstaw systemie nerwowym obejmuje rozumowanie. W takim zakresie logika jest nauką współtworzącą kognitywistykę.

W naukach ścisłych dominującą rolę odgrywa klasyczny rachunek logiczny jako podstawa pojęć i rozumowań matematycznych. Ważna jest również funkcja logicznej teorii rozumowań w metodologii tych nauk; umożliwia bowiem tworzenie modeli teorii naukowej oraz opisu procedur badawczych stosowanych w tych naukach. Z kolei logiki nieklasyczne dostarczają narzędzi dla formalizacji rozumowań języka wyobrażeniowego fizyki.

W naukach humanistycznych logiczne narzędzia analizy umożliwiają ujęcie metodologicznej specyfiki tych nauk (dominują w nich pojęcia typologiczne lub mające rodzinę znaczeń, a wyjaśnianie musi uwzględniać jednostkowość przedmiotu dociekań oraz zakładać racjonalność ludzkich podmiotów działających lub tworzących). Ważna jest rola logik nieklasycznych, dostarczających narzędzi analizy pojęć nauk humanistycznych oraz ontologii formalnych, korzystających z osiągnięć logiki klasycznej i logik nieklasycznych.

Szczególność logika odgrywa w naukach kognitywnych, dostarczając modeli ludzkich czynności poznawczych (np. logiki nieklasyczne, takie jak logiki dynamiczne, logiki wiedzy i przekonań, logiki działania, logiki deontyczne), umożliwiając formalne rekonstrukcje tych czynności poznawczych lub modeli umysłu.

Należy na koniec stwierdzić, że w praktyce zainteresowanie logikami nieklasycznymi po stronie fizyków okazało się niewielkie, natomiast takie aplikacje są rozpowszechnione po stronie filozofii nauk przyrodniczych.

Z kolei zastosowanie metod logicznych jest rozpowszechnione w metodologii nauk humanistycznych, a niektóre idee logiki (np. idea semantyk światów możliwych, ze względu na wielość i różnorodność filozoficznych interpretacji pojęcia światów możliwych) zyskały duży rozgłos wśród humanistów-nielogików.

BIBLIOGRAFIA

- Bocheński, I.M. (1961). *A History of Formal Logic*. Notre Dame: University of Notre Dame Press.
- Bocheński, I.M. (1993). *Logika i filozofia. Wybór pism*. Warszawa: PWN.
- Borkowski, L. (1970). *Logika formalna*. Warszawa: PWN.
- Grzegorzczak, A. (1981). *Zarys logiki matematycznej*. Warszawa: PWN.
- Hughes, G.E., & Cresswell, M.J. (1968). *An Introduction to Modal Logic*. London: Methuen Publ.
- Janeczek S., Starościc, A., & Tkaczyk, M. (Red.) (2018). *Logika. Cz. I: Natura logiki*. Lublin: Wydawnictwo KUL.
- Kamiński, S. (1992). *Nauka i metoda. Pojęcie nauki i klasyfikacja nauk*. Lublin: TN KUL.
- Kiczuk, S. (2001). *Przedmiot logiki formalnej oraz jej stosowalność*. Lublin: Red. Wyd. KUL.
- Kmita, J. (1971). *Z metodologicznych problemów interpretacji humanistycznej*. Warszawa: PWN.
- Lechniak, M. (2011). *Przekonania i zmiana przekonań*. Lublin: Wydawnictwo KUL.
- Łukasiewicz, J. (1961). *Z zagadnień logiki i filozofii*. Warszawa: PWN.
- Marciszewski, W. (Red.) (1987). *Logika formalna. Zarys encyklopedyczny*. Warszawa: PWN.
- Pawłowski, T. (1986). *Tworzenie pojęć w naukach humanistycznych*. Warszawa: PWN.
- Świrydowicz, K. (2004). *Podstawy logiki modalnej*. Poznań: Wyd. UAM.
- von Wright, G.H. (1951). *An Essay in Modal Logic*. Amsterdam: North Holland.

Anita Pacholik-Żuromska
Uniwersytet Mikołaja Kopernika w Toruniu
<https://orcid.org/0000-0002-8674-9412>
<https://doi.org/10.35765/slowniki.424>

Logiki niemonotoniczne w modelowaniu umysłu

Streszczenie

DEFINICJA POJĘCIA: Logiki niemonotoniczne odbiegają od klasycznych zasad wnioskowania, ponieważ nie zachowują zasady, według której jeśli zdanie wynika z pewnego określonego zbioru przesłanek, to wynika również z sumy tego zbioru z dowolnym zbiorem (tzw. zasada monotoniczności). Zadaniem logiki niemonotonicznej w modelach umysłu jest odzwierciedlenie sposobu ludzkiego myślenia. Stanowią one odpowiednie narzędzie do tworzenia formalnych modeli umysłu, ponieważ adekwatnie uchwytyją subtelności w potocznym rozumowaniu, takie jak pomijanie przesłanek czy wprowadzanie nowych założeń sprzecznych z poprzednimi.

ANALIZA HISTORYCZNA POJĘCIA: Od początku badań prowadzonych w ramach współczesnej filozofii umysłu, czyli mniej więcej od drugiej połowy XX wieku, modele umysłu oparte były na obliczeniowości. Jednak do opisu mechanizmów działania skomplikowanych procesów mentalnych potrzeba było czegoś więcej niż logiki klasycznej. Analogowe przetwarzanie informacji zastąpiono rozproszonym, a algorytmy wykorzystywane w symulacji ludzkich zdolności, jak uczenie się, stawały się coraz bardziej wyrafinowane. Obecnie renesans przeżywa kodowanie predykcyjne, które jednak często stosuje się w połączeniu z logikami niemonotonicznymi.

UJĘCIE PROBLEMOWE POJĘCIA: Logiki niemonotoniczne można wykorzystać do modelowania zarówno na poziomie przetwarzania

informacji neuronalnej, jak i formowania postaw propozycjonalnych. Z jednej strony bowiem niemonotoniczne wnioskowania w sieci neuronalnej są schematami zdefiniowanymi w terminach wektorów aktywności neuronalnej. Z drugiej strony mogą być traktowane jako struktura propozycjonalna, gdyż są powiązane niemonotoniczną relacją konsekwencji.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Zaletą korzystania z logik niemonotonicznych do modelowania umysłu jest to, że wspierają one tworzenie formalnych mechanizmów akumulacji wiedzy konceptualnej, czyli propozycjonalnej. Modele takie pokazują, w jaki sposób tworzona jest dynamiczna sieć przekonań podmiotu, np. w metapoznaniu.

Słowa kluczowe: model umysłu, logiki niemonotoniczne, reprezentacje, predykcje, niepewność

Definicja pojęcia i określenie obszaru badań

Logiki niemonotoniczne są rodzajem logik, w których wnioskowanie nie przebiega w sposób klasyczny, dla którego charakterystyczna jest zasada monotoniczności, głosząca, iż jeśli zdanie wynika z pewnego określonego zbioru przesłanek, to wynika również z sumy tego zbioru z dowolnym zbiorem (Makinson, 2008). Zatem wnioskowanie jest niemonotoniczne wtedy, gdy z pewnego zbioru Z wynika pewien wniosek p , który jednak nie wynika z pewnego nadzbioru tego zbioru (Makinson, 2008). W odróżnieniu od dedukcji inferencje niemonotoniczne powstają na bazie częściowej informacji (Moore, 1985), intuicji i oczekiwań (Gärdenfors & Makinson, 1994). Nowe informacje mogą prowadzić do zmiany wcześniejszych wniosków, ale przy zachowaniu poprzednich przesłanek i bez popadania w sprzeczność.

Z kolei modele umysłu stanowią zagadnienie interdyscyplinarne, angażujące między innymi filozofię, psychologię, logikę, neuronaukę czy lingwistykę. Innymi słowy, modele umysłu stanowią klasyczny problem dla kognitywistyki – interdyscyplinarnej (czy też transdyscyplinarnej) nauki o poznaniu. Przykładem praktycznego wykorzystania modelowania procesów poznawczych za pomocą odpowiednich narzędzi formalnych – algorytmów – jest zastosowanie sztucznej inteligencji w pojazdach autonomicznych.

Zadaniem logiki niemonotonicznej w modelach umysłu jest odzwierciedlenie sposobu przetwarzania informacji między innymi w procesie podejmowania decyzji, przede wszystkim jeśli chodzi o zdolność do elastycznego myślenia, uwzględnianie niepełnych informacji, zmianę perspektywy czy też żywienie sprzecznych przekonań. Chodzi o to, by taki model rzeczywiście pokazywał, jak działa ludzki umysł.

Modele umysłu stosujące logiki niemonotoniczne miały swoich poprzedników. Wcześniejsze modele również zawierały algorytmy, jednak oparte na logice klasycznej. Powstały jednak i tak dość późno, ponieważ w czasie, kiedy znane już były algorytmy stosowane w maszynach liczących i deszyfrujących. Jeszcze wcześniejsze modele umysłu odwoływały się natomiast do tego, co w danym czasie stanowiło aktualne osiągnięcie techniczne. Umysł porównywany był zazwyczaj do jakiegoś mechanizmu, czy to konstruowanych już w antycznej starożytnej Grecji pneumatycznych zabawek, czy zegara, którego działanie stanowiło inspirację dla rozważań Leibniza o umyśle.

Analiza historyczna pojęcia

Współczesna historia modeli umysłu zaczyna się w roku 1948. Wtedy to odbyło się pierwsze interdyscyplinarne sympozjum pt. „Cerebral Mechanisms and Behavior” (Mechanizmy mózgowe a zachowanie). Wśród uczestników byli między innymi John von Neumann, Paul Weiss i Warren McCulloch – pionierzy w badaniach nad sztuczną inteligencją. Pierwsze modele umysłu od początku nawiązywały zatem do paradygmatu obliczeniowego w kognitywistyce, opierającego się na teorii automatów, czyli matematycznych modeli urządzeń przetwarzających dyskretną informację. Dwa główne typy automatów to akceptujące (rozpoznające) i przetwarzające. Do ich zadań należy rozpoznawanie języków i obliczanie funkcji częściowych na zbiorze Σ^* , gdzie Σ i Δ to skończone alfabety, a Σ^* i Δ^* to zbiory wszystkich słów nad odpowiednimi alfabetami (por. Marciszewski, 1987, 200).

Od początku badań prowadzonych w ramach współczesnej filozofii umysłu (czy analitycznej metafizyki umysłu) i kognitywistyki, czyli mniej więcej od drugiej połowy XX wieku, powstały trzy główne modele oparte na obliczeniowości. Pierwszy z nich, posługujący się metaforą komputerową, głosił, że umysł jest rodzajem programu komputerowego – czy dosłownie komputera – a procesy w nim zachodzące mają charakter obliczeniowy: przebiegają na podstawie jakiegoś algorytmu. Modelem umysłu, który powstał na podstawie tego założenia, była Reprezentacyjna Teoria Umysłu autorstwa Jerry’ego A. Fodora. Zgodnie z nią treść stanów mentalnych determinowana jest przez wewnętrzne relacje zachodzące między reprezentacjami mentalnymi, będącymi nośnikiem informacji. To na reprezentacjach (tzw. symbolach języka myśli) przeprowadzane są obliczenia (Fodor, 1999). Fodor rozumiał obliczeniowość dość prosto, jako zgodną z zasadami logiki klasycznej (klasycznego rachunku zdań i klasycznego rachunku predykatów) i w ten sposób wyjaśniał racjonalne rozumowanie podmiotu.

Jasne jednak było, że umysł jest bardziej skomplikowany i logika klasyczna nie wystarczy do opisu mechanizmów działania skomplikowanych procesów mentalnych. Weźmy np. zjawisko akrazji, czyli słabości woli, w którym podmiot żywi sprzeczne przekonania na temat tego, co powinien i co chce uczynić, np., że palenie szkodzi zdrowiu, zatem powinien natychmiast rzucić palenie, ale z drugiej strony nie ma ochoty

rzucac palenia. W każdym razie modele rozumowań, w których podmiot przechodzi od jednego prawdziwego przekonania do drugiego (Fodor, 1999), były zbyt wyidealizowane. Zaczęto zatem stosować metaforę mózgu, odwracając perspektywę. To nie komputer miał stanowić model dla naturalnego umysłu, ale działanie mózgu, symulowane na sztucznych sieciach neuronowych miało wyjaśnić, jak działa umysł. Komputerową metaforę umysłu zastąpił koneksjonizm. Koneksjonizm jako formalne narzędzie modelowania umysłu stosował już nie szeregowe przetwarzanie informacji, ale rozproszone, dla którego model sieci był najbardziej odpowiedni, ponieważ „prawie każda część mózgu jest obserwowana przez inną, a często przez wiele innych naraz” (Churchland, 2002, s. 17). Zaczęto zatem budować sztuczne sieci neuronowe (SSN), które symulowały np. takie procesy mentalne jak rozpoznawanie twarzy. W architekturze koneksjonistycznej przetwarzanie informacji przebiega w ten sposób, że wzorce aktywacji neuronów w narządach zmysłowych monitorowane są przez kolejne warstwy neuronów, przekształcające daną informację i przekazujące ją do kolejnych struktur w mózgu (Churchland, 2002). Informacja zakodowana w mózgu to wzorzec aktywacji neuronalnej, który może być wzmacniany na przykład przez trening poznawczy. Modele koneksjonistyczne pozwalają na zdefiniowanie procedur uczenia się, dzięki czemu dopasowują się do środowiska.

Jednak modele koneksjonistyczne, choć aplikowane na SSN, nie spełniały całkowicie funkcji wyjaśnienia, jak działa umysł, między innymi dlatego, że w nieznacznym stopniu uwzględniały udział ciała i otoczenia podmiotu w poznaniu. Modele, które miały to zmienić, zaczęły powstawać w ramach tzw. teorii 4-E, według których umysł jest ucieleśniony (*embodied*), rozszerzony (*extended*), zakorzeniony, czyli usytuowany w środowisku (*embedded*), a poznanie ma charakter enaktywny, tzn. zachodzi w bezpośredniej interakcji podmiotu z otoczeniem. Podkreślając rolę ciała w poznaniu, teorie 4-E nawiązywały do tradycji fenomenologicznej. W tym kontekście pojawiło się pytanie, jak modelować ten rodzaj poznania, w które zaangażowany jest nie tylko mózg, ale cały organizm wraz z jego otoczeniem. Najlepiej pasujące wydawały się wyjaśnienia posługujące się statystyką i prawdopodobieństwem. Problemem było jednak to, że teorie korzystające z takich obliczeń spolaryzowały dyskusję, budząc dawny spór internalizmu z eksternalizmem w filozofii

umysłu, co stanowiło nową odsłonę dawnej debaty m.in. między Fodorem a Putnamem (Fodor, 1999). Spór ten był kontynuacją jeszcze wcześniejszego problemu dotyczącego tego, jak przebiega poznanie: czy od umysłu do świata, czy też od świata do umysłu. Według założeń teorii 4-E nie tylko poznanie i znaczenie miałyby być rozszerzone, ale także reprezentacje, ponieważ bodziec ze świata jest elementem reprezentacji mentalnej, a nie tylko jej aktywatorem.

Charakterystyczne dla modeli w teoriach 4-E było po pierwsze uwzględnienie informacji zwrotnej z otoczenia (*feedback*), po drugie nałożenie na reprezentacje wymogu dynamiczności lub też całkowite ich odrzucenie. Badacze, aby uniknąć konieczności budowania modeli na podstawie pojęcia reprezentacji, zwrócili się w kierunku cybernetyki, zapożyczając z niej np. pojęcie niezmiennika zastępującego rolę reprezentacji w pętli interakcji mózg – ciało – świat (Seth, 2015, 1465).

Elementem unifikującym dla modeli uwzględniających *feedback*, zachowujących reprezentacje i jednocześnie odwołujących się do działania mózgu – czyli w pewnym sensie jednoczących koneksjonizm z teoriami 4E – okazało się kodowanie predykcyjne. Model obliczeniowy oparty na Bayesowskiej teorii indukcji w połączeniu z zasadą swobodnej energii (Friston, 2009) proponuje mechanistyczne wyjaśnianie funkcjonowania umysłu. Zgodnie z wnioskowaniem Bayesowskim prawdopodobieństwo danej hipotezy jest aktualizowane przy większym dostępie do danych – nowej informacji. Jest to prawdopodobieństwo *a posteriori*, które oddaje potoczną intuicję, że „wynik obserwacji potwierdza hipotezę wówczas, gdy jest wysoce prawdopodobny na jej gruncie, a poza tym – mało prawdopodobny” (Marciszewski, 1987, s. 412). Bayesowski mózg działa w ten sposób, że doświadczenie (*feedback* ze świata) jest konfrontowane z uprzednią predykcją (hipotezą) na temat spodziewanej sytuacji. Owe hipotezy można interpretować jako reprezentacje, będące modelem prawdopodobnego (oczekiwanego) stanu świata. Kodowanie predykcyjne stosuje się na przykład w formalizacji nienadzorowanego uczenia sztucznych sieci neuronowych, które jest używane do samoorganizacji systemu zarówno w odniesieniu do jego elementów, jak i środowiska. Ten mechanizm został przeniesiony na naturalny system poznawczy, aby wyjaśnić mechanizmy poznawcze organizmów żywych. W tej koncepcji samoorganizacja to ciągły proces dopasowywania reprezentacji umysłowych do świata. Celem takiego mechanizmu

jest minimalizacja błędu predykcyjnego, czyli optymalizacja działania. Dwa odrębne, ale skorelowane aspekty fundamentalnego mechanizmu mózgu, który minimalizuje błędy predykcyjne, to percepcja i uwaga. Percepcje są wynikiem neuronalnego modelu świata (modelu odzwierciedlającego, jak mózg „widzi” rzeczywistość), w którym oczekiwania i przewidywania dotyczące stanów świata są porównywane z bodźcem ze świata. Choć wnioskowanie Bayesowskie może radzić sobie z modyfikacją wcześniej przyjętych przekonań, to napotyka trudności w kontekstach, gdzie podmiot nie posiada pełnej wiedzy o świecie. Dlatego badacze poszukują jeszcze lepszych narzędzi formalnych, nie wykluczając jednak i nie przekreślając kodowania predykcyjnego. Jedną z propozycji są modele procesów przetwarzania informacji, oparte na logikach niemonotonicznych.

Ujęcie problemowe pojęcia

W pytaniu o to, czy ludzkie rozumowanie jest niemonotoniczne, badacze skłaniają się ku pozytywnej odpowiedzi (Malinowski, 1997). Wnioskowanie Bayesowskie, choć radzi sobie z modyfikacją wcześniej przyjętych przekonań, napotyka trudności w kontekstach, gdzie podmiot nie ma pełnej wiedzy o świecie. Mówimy tu o kontekstach odpowiadających realnym codziennym sytuacjom (Mota, Sridharan & Leonardis, 2021). Również w projektowaniu interakcji człowiek – robot przetwarzanie predykcyjne występuje w połączeniu z logikami niemonotonicznymi, co daje lepsze efekty na przykład w programowaniu procesu podejmowania odpowiednich decyzji i szybkiego, a przy tym trafnego reagowania (Riley & Sridharan, 2019).

Logiki niemonotoniczne są formalnymi strukturami używanymi do uchwycenia tzw. *defeasible reasoning*, czyli typowego ludzkiego rozumowania, które jest racjonalne, ale nie jest dedukcyjne. W przeciwieństwie do dedukcji inferencje niemonotoniczne są przeprowadzane na podstawie częściowej informacji (Moore, 1985), intuicji i oczekiwań (Gärdenfors & Makinson, 1994). Co więcej, możemy dojść do nowego wniosku, gdy otrzymamy nowe informacje, nie odrzucając żadnych wcześniejszych przesłanek i bez wprowadzania sprzeczności. Konkretnie rozszerzenia, takie jak dodawanie nowych przesłanek, sprawiają, że

logika staje się niemonotoniczna. Jest to dobrze znane w potocznym rozumowaniu:

*Zazwyczaj ptaki latają.
Struś Pędziwiatr to ptak.
Zatem Struś Pędziwiatr lata.*

Wniosek w tym wypadku jest jednak fałszywy, ponieważ strusie to nie-loty. Niemonotoniczne reguły wnioskowania pozwalają nam rozwiązać błąd powstały na skutek generalizacji:

Struś Pędziwiatr to ptak, więc Struś Pędziwiatr może latać.
 $\{p\} \vdash z$
Struś Pędziwiatr to ptak. Struś Pędziwiatr to struś, więc Struś Pędziwiatr nie może latać.
 $\{p, q\} \nvdash z$

Klasyczna relacja konsekwencji (\vdash) jest silniejsza niż relacja niemonotoniczna \vdash . Jako operacja, klasyczna konsekwencja (C_n) to operacja domknięcia (Makinson 2005, 4). Oznacza to, że spełnia ona następujące warunki, które są charakterystyczne dla wnioskowania dedukcyjnego (Malinowski, 1997):

Zwrotność (inkluzja): $A \subseteq C_n(A)$
 (Zbiór A jest podzbiorem zbioru wszystkich konsekwencji zbioru A).
 Idempotentność: $C_n C_n(A) = C_n(A)$
 (Zbiór konsekwencji logicznej konsekwencji przesłanek jest równy zbiorowi konsekwencji przesłanek).
 Monotoniczność: Zawsze, gdy $A \subseteq B$, to $C_n(A) \subseteq C_n(B)$
 (Konsekwencja danego zbioru przesłanek jest także konsekwencją dowolnego rozszerzenia tego zbioru).

Logika niemonotoniczna jest natomiast logiką, w której warunek monotoniczności nie jest spełniony, to znaczy nie jest tak, że jeśli zdanie wynika z pewnego określonego zbioru przesłanek, to wynika również z sumy tego zbioru z dowolnym zbiorem:

$$x \in C_n(A) \wedge x \notin A \cup B$$

gdzie x jest konsekwencją zbioru A zdań, ale nie jest konsekwencją nadzbioru $A \cup B$ tego zbioru (Makinson, 2005, s. 2; 2005a, s. 4).

Ale to właśnie słabsza relacja konsekwencji towarzyszy zwykłemu, codziennemu rozumowaniu, które tylko czasami jest dedukcyjne. Ludzie nie są maszynami Turinga, a ich sądy często opierają się na pewnego rodzaju pomieszaniu dedukcji i indukcji, co jest charakterystyczne dla niemonotoniczności. Podczas gdy klasyczna konsekwencja dobrze „radzi sobie” z dedukcyjnym rozumowaniami, gdzie mamy do czynienia z semantyką i syntaktyką, zawodzi tam, gdzie zaangażowana jest pragmatyka, jak to ma miejsce w przypadku aktów komunikacyjnych, czyli tam, gdzie w grę wchodzi intencje, przekonania, czy generalnie elementy psychologiczne.

Logiki niemonotoniczne już udowodniły swoją skuteczność w modelach konkretnych działań poznawczych, implementowanych na sztucznych systemach, m.in. w oprogramowaniu dla robotów (Riley & Sridharan, 2019; Mota, Sridharan & Leonardis, 2021). Te architektury używają logik niemonotonicznych do podejmowania decyzji, przewidywania zachowań, modelowania procesów uczenia, programowania odpowiednich reakcji, przekonań i doświadczeń (Riley & Sridharan, 2019). Zaletą korzystania z logik niemonotonicznych do modelowania umysłu jest to, że wspierają one tworzenie formalnych mechanizmów akumulacji wiedzy konceptualnej, czyli propozycjonalnej. Model formowania takiej wiedzy, skonstruowany na podstawie logik niemonotonicznych, pokazuje, jak rozszerzenie zbioru przesłanek, czyli dodanie nowej wiedzy, zmienia również jej konsekwencję. Wnioski są zatem wyciągane na podstawie dostępnej informacji w taki sposób, że mogą być odrzucone, jeśli pojawią się dodatkowe informacje, ale bez odrzucania którejkolwiek ze starych przesłanek (por. Makinson, 2005; Makinson, 1994).

Logiki niemonotoniczne można wykorzystać do modelowania zarówno na poziomie przetwarzania informacji neuronalnej, jak i formowania postaw propozycjonalnych (Balkenius & Gärdenfors, 1991). Na poziomie intencjonalnym, czyli propozycjonalnym, logiki niemonotoniczne mogą modelować wymianę przekonań w konwersacji, ponieważ obejmują symboliczny i subsymboliczny poziom przetwarzania informacji, co oznacza, że opisują jednocześnie wektory aktywności neuronalnej i językową formę przekonań (Balkenius & Gärdenfors, 1991, s. 37; Blutner, 2004). Jest to możliwe, ponieważ – chociaż niemonotoniczne wnioskowania w sieci neuronalnej są schematami zdefiniowanymi w terminach wektorów aktywności neuronalnej – mogą być traktowane jako

struktura propozycjonalna, gdyż są powiązane niemonotoniczną relacją konsekwencji \vdash (Balkenius & Gärdenfors, 1991, s. 36). Logiki niemonotoniczne wypełniają lukę między dwoma różnymi trybami obliczeń: modelem symbolicznym (jak np. w Reprezentacyjnej Teorii Umysłu Fodora) i subsymbolicznym (neuronalnym). Mogą być zatem stosowane w modelowaniu zachowania sztucznych sieci neuronowych, ponieważ działania sieci koneksjonistycznych można interpretować jako niemonotoniczne wnioskowania. Innymi słowy: logiki niemonotoniczne można wykorzystać jako narzędzie opisowe i analityczne do analizy emergentnych właściwości sieci koneksjonistycznych.

Balkenius i Gärdenfors (1991) udowodnili, że możliwe jest zdefiniowanie relacji niemonotonicznego wnioskowania, która odzwierciedla asymptotyczne aktualizacje stanów informacyjnych;

Niech $\langle S, \geq \rangle$ będzie częściowo uporządkowanym zbiorem stanów aktywacji, a w macierzą połączeń. Wówczas pojęcie asymptotycznych aktualizacji naturalnie prowadzi do niemonotonicznej relacji wnioskowania między stanami informacyjnymi:

$$s \vdash_w t \text{ jeżeli } s' \geq t \text{ dla każdego } s' \in \text{ASUP}(s)$$

Logiki niemonotoniczne charakteryzują asymptotycznie, w jaki sposób aktywność neuronów rozprzestrzenia się przez sieć koneksjonistyczną. Korespondencja między macierzą połączeń a systemem Poole'a (1988) z adnotacjami wagowymi polega na zastosowaniu mechanizmu translacji, w którym węzeł i jest przekształcany w symbol atomowy p_i . Aktywujące połączenie w sieci jest przekształcane w dwukierunkowy logiczny równoważnik \leftrightarrow , a hamujące połączenie w równoważnik \leftrightarrow z negacją \sim jednego z jego argumentów. Wagi domyślne należy traktować jako wartość bezwzględną odpowiadających im elementów macierzy (Bluntner, 2004). Przy użyciu logik niemonotonicznych zachodzi ścisła zgodność między reprezentacją wiedzy w systemie Poole'a (używanego w różnych dziedzinach sztucznej inteligencji i korzystającego z logiki założeń domyślnych, jak w przykładzie z przesłanką: „Ptaki latają”) a kodowaniem wiedzy w sieciach Hopfielda, tj. sieciach rekurencyjnych stosowanych np. w koneksjonistycznych modelach pamięci asocjacyjnej.

Jak zostało powiedziane powyżej, choć kodowanie predykcyjne stanowi bardzo dobre narzędzie formalne do modelowania funkcji

i zdolności poznawczych, to jednak w wypadku ludzkiego rozumowania nie wystarcza. Z jednej strony, formalizacja za pomocą kodowania predykcyjnego dokładnie opisuje związek między wewnętrzną reprezentacją a światem zewnętrznym, ponieważ wyjaśnia, jak mechanizm korygowania reprezentacji i dostosowywania zachowania do otoczenia działa na poziomie neuronalnym (Chanes & Barrett, 2016). Należy również podkreślić, że kodowanie predykcyjne jest używane do modelowania sytuacji konwersacyjnych, na przykład do ustalania kolejności wypowiedzi, co można wytłumaczyć funkcjonalnym mechanizmem przewidywania umożliwiającym zmienianie roli z mówcy na słuchacza (Riest, Jorschick & de Ruiter, 2015). Z drugiej strony logiki niemonotoniczne nadają się do modelowania predykcji na podstawie stawianych hipotez czy też oczekiwań dotyczących przyszłego stanu otoczenia (Gärdenfors & Makinson, 1994). Problem polega jednak nie na tym, jak uchwycić przewidywalność, ale jak poradzić sobie z niepewnością (Talamadupula et al., 2010). Termin „niepewność” często jest używany nieprawidłowo jako odniesienie do sytuacji, gdy nie znamy dokładnie tego, co się stanie, gdy podejmiemy działania lub wprowadzimy inne zmiany w dziedzinie. W rzeczywistości jest to tak zwana niejednoznaczność, która w ekonomii i teorii prawdopodobieństwa znana jest także jako ryzyko. Niepewność występuje, gdy nie znamy wszystkich alternatyw, wyników lub prawdopodobieństw. Świat, w którym nie możemy znać wszystkich tych możliwości, to „otwarty świat” (Open World), którego ontologia obejmuje wiele realnych problemów wynikających z niepewności. „W zamkniętym świecie to, co nie jest znane jako prawdziwe, musi być fałszywe. W otwartym świecie to, co nie jest znane jako prawdziwe, po prostu jest nieznanie” (Sequeda 2012). Dlatego potrzebujemy właściwej semantyki otwartego świata do reprezentowania treści takiej częściowej wiedzy. W tym przedsięwzięciu badacze używają heurystyk, ale nie stosują poprawnych formalnych modeli do projektowania heurystycznego. Choć poświęcono już wiele czasu na próby dostosowania metod wnioskowania Bayesowskiego do radzenia sobie z niepewnością (Riley & Sridharan, 2019), ta metoda nie jest do tego przeznaczona. „Modelowanie niepewności przez dodawanie prawdopodobieństw do twierdzeń logicznych nie zawsze ma sens” (Riley & Sridharan 2019, s. 4). W związku z tym, jeśli chcemy stworzyć modele działania używane do wnioskowań czy też uczenia się w dynamicznym środowisku,

należy połączyć wnioskowania niemonotoniczne z relacyjnym uczeniem się ze wzmocnieniem (*reinforcement learning*) oraz uczeniem indukcyjnym (Riley & Sridharan 2019, s. 4).

Logiki niemonotoniczne pozwalają na kwestionowanie wcześniej przyjętych założeń ze względu na swoją naturę. Mają także większy potencjał do kodowania i przeprowadzania tego rodzaju rozumowania wymaganego do rzeczywistego radzenia sobie z niepewnością przy pewnym treningu (Mota, Sridharan & Leonardis, 2021). Kodowanie predykcyjne oparte na wnioskowaniu Bayesowskim nie wystarcza do stworzenia odpowiedniej architektury opisującej ludzkie rozumowanie, nie tylko dlatego, że nie radzi sobie z niepewnością, ale także dlatego, że jest problematyczne w kontekście indukcji (Kyburg, 2001). Z kolei logiki niemonotoniczne sprawdzają się w kodowaniu złożonych wyrażeń i subtelnych różnic, takich jak różnica między „ $\neg a$ ” (co oznacza, że „ a jest uważane za fałszywe”) a „nie a ” (co tylko oznacza, że „nie uważa się, że a jest prawdziwe”) (por. Mota, Sridharan & Leonardis, 2021, s. 242). Dzięki temu możliwe jest rozróżnienie między negacją a alternatywą epistemiczną. Natomiast dodanie nowych przesłanek pozwala na zmniejszenie zbioru konsekwencji, a przez to szybsze naprawienie błędów inferencyjnych (Mota, Sridharan & Leonardis, 2021, s. 242).

Dyskusja i refleksja

Przykłady praktycznych zastosowań logik niemonotonicznych w sztucznych sieciach neuronowych, czyli ich wzajemnej integracji, obejmują rozwinięcie systemów zdolnych do niezawodnego i efektywnego rozpoznawania oraz reagowania np. na różne znaki drogowe (Riley & Sridharan, 2019), co jest istotne choćby w programowaniu autonomicznych pojazdów. Z kolei próba przeniesienia mechanizmu działania sztucznych sieci neuronowych na mechanizm umysłu wpisuje się w tradycję koneksjonizmu.

Biorąc pod uwagę to, że logiki niemonotoniczne znakomicie nadają się do modelowania wiedzy propozycjonalnej, nasuwa się wniosek, że stanowią one też bardzo dobre narzędzie w modelowaniu metapoznania, szczególnie tego w postaci propozycjonalnej, a więc sądów samowiedzy rozumianej jako pierwszoosobowa wiedza podmiotu o własnych

stanach mentalnych, do których ma on uprzywilejowany dostęp. Są to np. sądy typu: „Jestem przekonany, że jestem przekonany, że 21 marca zaczyna się kalendarzowa wiosna”. Innymi słowy są to pierwszoosobowe sprawozdania z własnych przekonań, pragnień, decyzji itd. – czyli przekonania o przekonaniach. W tym sensie samowiedza jest reprezentacją przekonań drugiego rzędu, które tworzą wiedzę podmiotu na temat treści jego umysłu.

Badacze zaczęli ponownie rozważać pomysł modelowania konkretnych działań poznawczych przy użyciu logik niemonotonicznych, ponieważ chcieli zaimplementować w robotach rodzaj samowiedzy, czyli w tym wypadku „zdolność do dostarczania wyjaśniających opisów swoich decyzji” (Mota, Sridharan & Leonardis, 2021, s. 241). Chociaż wnioskowania o świecie są Bayesowskie, tj. przetwarzanie predykcyjne może być narzędziem do modelowania przekonań pierwszego rzędu, przekonania drugiego rzędu (czyli przekonania o przekonaniach, które są charakterystyczne dla samowiedzy) muszą być modelowane za pomocą logik niemonotonicznych, ponieważ „umożliwiają naprawę błędów wynikających z błędnych inferencji” (Mota, Sridharan & Leonardis, 2021, s. 242–243). Elementy probabilistyczne nie pozwalają na przeprowadzanie skutecznych wnioskowań w ramach wiedzy zdroworozsądkowej (Riley & Sridharan, 2019).

Zastosowanie logik niemonotonicznych do modelowania samowiedzy nie jest niczym nowym. Zostały one użyte w logice autoepistemicznej do stworzenia modelu idealnie racjonalnego podmiotu rozumującego o własnych przekonaniach (Moore, 1985). Logiki niemonotoniczne używane do reprezentacji wiedzy mogą być zatem również zastosowane w reprezentacji samowiedzy. Ten rodzaj wiedzy wydaje się bowiem szczególnym przypadkiem, w którym podmiot jest w stanie utrzymać sprzeczne przekonania, jak np. to, że palenie szkodzi zdrowiu, i to, że nie ma ochoty rzucić palenia. Problem niekoherentnych przekonań w samowiedzy wiąże się ze sposobem determinacji ich treści. Wpływ na to, jak podmiot postrzega siebie, zależy od zewnętrznych czynników pochodzących ze społecznego otoczenia podmiotu. Zmiana otoczenia lub grupy, do której należy podmiot, może mieć głęboki wpływ na jego przekonania. Czasem wskutek takiej zmiany osoba może przyjąć przekonania, o których wcześniej nie miała pojęcia, co skutkuje odmiennym spojrzeniem na samą siebie. To zjawisko samo w sobie nie jest

niczym nowym, ponieważ ludzie podlegają zmianom, dostosowując się do nowych sytuacji życiowych. Interesujące jest pytanie, jak mimo tych zmian zachowujemy spójność psychiczną i poczucie tożsamości. W trakcie życia dostosowujemy się do nowych warunków i wyzwań. Aby nie stracić poczucia „bycia sobą”, kształtujemy naszą własną narrację zgodnie z nowymi przekonaniem, które mogły ulec zmianie w wyniku różnych czynników. Zmiany w naszych przekonaniach wpływają na naszą percepcję samych siebie. Staramy się dopasować kolejne przekonania w taki sposób, aby zachować spójność, choć nie zawsze jesteśmy w stanie precyzyjnie uzasadnić, dlaczego akceptujemy jedno przekonanie, a odrzucamy inne.

Metoda unifikacji pewnych formalizmów niemonotonicznych została zaproponowana przez Przymusińskiego (1997). Formalna maszyna oferowana jako alternatywa dla kodowania predykcyjnego składa się z połączenia logik niemonotonicznych z drzewem decyzyjnym. Jest to metoda uczenia indukcyjnego, do której zalicza się uczenie ze wzmocnieniem. W przypadku sprzecznych przekonań i paradoksów logiki niemonotoniczne stosują kryterium praktyczności, jako składnika racjonalności (por. Makinson, 2005, s. 129–130). Wykorzystanie logik niemonotonicznych do architektury umysłu, rozumianego w kategoriach postaw propozycjonalnych, pokazuje, jak tworzona jest dynamiczna sieć przekonań podmiotu, gdzie zmiana jednego przekonania prowadzi do zmiany innych przekonań, a dodanie nowego przekonania może potencjalnie wpłynąć na całą sieć.

Językowa forma samowiedzy wyrażona w formie raportów pierwszoosobowych, takich jak „Jestem przekonany, że jestem przekonany, że p ”, składa się z dwóch podstawowych reprezentacji: reprezentacji Ja i reprezentacji doświadczanego stanu, połączonych relacją intencjonalności wyrażającą postawę podmiotu wobec danego sądu o sobie. Ze względu na swój charakter reprezentacyjny samowiedza może być modelowana obliczeniowo (por. Lewis, 2015) i implementowana na sztucznych sieciach neuronowych przy użyciu logik niemonotonicznych.

Ale logiki niemonotoniczne mogą również być dobrym narzędziem do modelowania zaburzeń psychicznych i radzenia sobie z auto-iluzjami oraz fałszywymi przekonaniem, na przykład w odniesieniu do refleksji podmiotu na temat swoich własnych przekonań, gdy odkrywa, że są one sprzeczne lub „dziwaczne” (Moore, 1985). Dotyczy

to na przykład pacjentów cierpiących na schizofrenię, którzy relacjonując swoje doświadczenia podczas stanu psychotycznego, oświadczają, że wiedzą, iż to była halucynacja, ale jednocześnie wierzą, że dane zdarzenie faktycznie miało miejsce. Innymi słowy są w pewien sposób odporni na zawodność swoich wnioskowań. To zjawisko można wytłumaczyć w ramach niemonotoniczności pozostawiającej otwartą możliwość, że podmiot może wierzyć w twierdzenia, które w żaden sposób nie są oparte na ich początkowych przesłankach (Moore, 1985). Innymi słowy modele samowiedzy oparte na logikach niemonotonicznych wyjaśniają, dlaczego podmiot może się mylić co do swoich własnych stanów umysłowych, jednocześnie utrzymując spójność przekonań tworzących samowiedzę

Jeśli rozważamy samowiedzę jako wewnętrzną mowę ze względu na jej treść sformułowaną w postaci sądów, to tym bardziej nadaje się ona do formalizacji za pomocą logik niemonotonicznych. Ponieważ logiki niemonotoniczne nie podlegają regule substytucji w swoich właściwych, ponadklasowych częściach, umożliwiają wyciąganie wniosków, które wykraczają poza zakres logiki klasycznej (por. Jarmużek, 2007). Dlatego też pozwalają na rozpoznawanie pozornie sprzecznych sądów i ciągłe aktualizowanie samowiedzy. „Ta właściwość różni się od logiki klasycznej, gdzie prawdziwe wnioski pozostają niezmiennie w obliczu nowych informacji” (Jarmużek, 2007, s. 259). Ważne jest tutaj to, że brakujące informacje są otrzymywane „z zewnątrz” w procesie dopasowywania siebie do innych podczas interakcji. Zatem w modelach umysłu i poznania konieczne jest wyjście poza paradygmat kodowania predykcyjnego, który wspiera raczej internalistyczne modele umysłu. Jest to szczególnie istotne, jeśli te rozważania mają obejmować kwestię sztucznej inteligencji.

Podsumowując, użycie logik niemonotonicznych w odniesieniu do postaw propozycyjalnych, czyli architektury umysłu rozumianej w kategoriach „psychologii ludowej” (Fodor, 1999), pokazuje, w jaki sposób tworzona jest dynamiczna sieć przekonań podmiotu, gdzie zmiana jednego przekonania prowadzi do zmiany innych przekonań, a dodanie nowego przekonania może potencjalnie wpłynąć na całą sieć. Logiki niemonotoniczne stanowią odpowiednie narzędzie do tworzenia formalnych modeli umysłu, między innymi metapoznania, ponieważ adekwatnie uchwytują subtelności, takie jak pomijanie przesłanek czy

wprowadzanie nowych założeń sprzecznych z poprzednimi, co też faktycznie ma miejsce w procesie zmiany przekonań. Inferencje niemonotoniczne odzwierciedlają, jak ludzie rozumują, decydują i weryfikują stan swojej wiedzy. Z perspektywy interakcji człowiek – robot, modele oparte na logikach niemonotonicznych pozwalają sztucznym systemom radzić sobie z niepewnością w tzw. otwartym świecie, gdzie wiele czynników po prostu jest nieznanych ze względu na niedostateczną informację.

BIBLIOGRAFIA

- Balkenius, C., & Gärdenfors, P. (1991). Nonmonotonic Inferences in Neural Networks. W: J. Allen, R. Fikes, & E. Sandewall (Red.), *Principles of Knowledge Representation and Reasoning. Proceedings of the Second International Conference (KR91)* (s. 32–39). California: Morgan Kaufmann Publishers.
- Blutner, R. (2004). Nonmonotonic Inferences and Neural Networks. W: W. van der Hoek (Red.), *Information, Interaction and Agency* (s. 203–234). Berlin: Springer, Dordrecht. DOI: 10.1007/1-4020-4094-6_7.
- Churchland, P.M. (2002). *Mechanizm rozumu, siedlisko duszy. Filozoficzna podróż w głąb mózgu*, Warszawa: Aletheia.
- Fodor, J.A. (1999). *Jak grać w reprezentacje umysłowe – poradnik Fodora*. Przeł. A. Putko. W: Z. Chlewiński (Red.), *Modele umysłu*. Warszawa: PWN.
- Friston, K.J. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293–301. DOI: 10.1016/j.tics.2009.04.005.
- Gärdenfors, P., & Makinson, D. (1994). Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65(2), 197–245. DOI: 10.1016/0004-3702(94)90017-5.
- Jarmużek, T. (2007). Book Review: David Makinson, “Bridges from Classical to Nonmonotonic Logic”, King’s College Publications, London, 2005, pp. 216, ISBN 1-904987-00-1. *Logic and Logical Philosophy*, 16, 259–262. DOI: 10.12775/LLP.2007.007.
- Makinson, D. (2008). *Od logiki klasycznej do niemonotonicznej*. Przeł. T. Jarmużek. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Makinson, D. (2005). Bridges from Classical to Nonmonotonic Logic. *Text in Computing*, 5. London: King’s College Publications.
- Malinowski, J., (1997). Logiki niemonotoniczne. Czy logika jest nauką kognitywną? *Przegląd Filozoficzny – Nowa seria*, R. VI, 1(21), 31–53.

- Marciszewski, W. (Red.) (1987). *Logika formalna. Zarys encyklopedyczny z zastosowaniem do lingwistyki i informatyki*. Warszawa: PWN.
- Moore, R.C., (1985). Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25(1), 75–94.
- Mota, T., Sridharan, M., & Leonardis, A. (2021). Integrated Common-sense Reasoning and Deep Learning for Transparent Decision Making in Robotics. *SN Computer Science*, 2(4), 242. DOI: 10.1007/s42979-021-00573-0.
- Poole, D. (1988). A logical framework for default reasoning. *Artificial Intelligence*, 36(1), 27–47.
- Przymusiński, T.C. (1997). Autoepistemic logic of knowledge and beliefs. *Artificial Intelligence*, 95(1), 115–154. DOI: 10.1016/S0004-3702(97)00032-5.
- Riley, H., & Sridharan, M. (2019). Integrating Non-monotonic Logical Reasoning and Inductive Learning With Deep Learning for Explainable Visual Question Answering. *Frontiers in Robotics and AI*, 6(125). DOI: 10.3389/frobt.2019.00125.
- Seth, A.K. (2015). The Cybernetic Bayesian Brain – From Interoceptive Inference to Sensorimotor Contingencies. W: T. Metzinger, & J.M. Windt (Red.). *Open MIND*, 35, 1–24. DOI: 10.15502/9783958570108.
- Sequeda, J. (2012, 30 listopada). *Introduction to: Open World Assumption vs Closed World Assumption, Dataversity*. Pobrano z: <https://www.dataversity.net/introduction-to-open-world-assumption-vs-closed-world-assumption/#> (dostęp: 30.11.2022).

Andrzej Bielecki
AGH Akademia Górniczo-Hutnicza
<https://orcid.org/0000-0002-0192-3785>
<https://doi.org/10.35765/slowniki.417>

Cybernetyka

Streszczenie

DEFINICJA POJĘCIA: Cybernetyka jest nauką interdyscyplinarną o sterowaniu w systemach, które działają w pewnym środowisku. System poprzez swoje moduły wejściowe odbiera ze środowiska sygnały oraz – w przypadku niektórych typów systemów – pobiera energię i materię, a także oddziałuje na środowisko przy pomocy swoich modułów wyjścia. U podstaw cybernetyki leży założenie, że niezależnie od tego, czy system jest maszyną, organizmem żywym czy strukturą społeczną, procesy związane ze sterowaniem i, co za tym idzie, przesyłaniem i przetwarzaniem w systemie sygnałów, informacji i energii w kontekście sterowania, mają wspólne cechy. Szczególną uwagę poświęca się w cybernetyce systemom autonomicznym, czyli takim, które same wyznaczają sobie cele, opracowują algorytmy ich realizacji oraz zapewniają sobie stabilność funkcjonalną.

ANALIZA HISTORYCZNA POJĘCIA: Rozwój cybernetyki jest ściśle związany z rozwojem matematyki, informatyki, robotyki i biologii, gdyż cybernetyka posiada ze wszystkimi tymi naukami wspólne i niezwykle ważne obszary badawcze. Począwszy od lat 30. XX wieku powstały teoretyczne podstawy cybernetyki stworzone przez matematyków (koncepcja maszyny Turinga) i lekarzy (pojęcie homeostazy). Istotny wpływ na cybernetykę miał rozwój techniki obliczeniowej – wprowadzenie komputerów i ich systematyczne ulepszanie. Współcześnie główne nurty badań cybernetycznych związane są ze sztuczną inteligencją, robotyką i cybernetyką układów biologicznych.

UJĘCIE PROBLEMOWE POJĘCIA: Ze względu na uniwersalność narzędzi badawczych, m.in. uniwersalność koncepcji systemu cybernetycznego,

cybernetyka jest obecnie stosowana w różnych dziedzinach. Postępując się metodami cybernetycznymi, analizuje się systemy ekonomiczne, społeczne, techniczne, biologiczne, polityczne, a nawet kulturowe. Cybernetyka dostarcza metod do analizy struktury i dynamiki systemu niezależnie od jego nośnika, stąd możliwości zastosowań w tak różnorodnych obszarach.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Cybernetyka wnosi istotny wkład do modelowania i analizy złożonych zjawisk. Zarówno w aspekcie historycznym, jak i współcześnie polscy naukowcy istotnie wpływają na rozwój cybernetyki.

Słowa kluczowe: system cybernetyczny (agent), sterowanie, informacja, homeostaza, sprzężenie zwrotne

Definicja pojęcia i określenie obszaru badań

Cybernetyka (gr. κυβερνήτης – sternik) jest teoretyczną, interdyscyplinarną nauką o przetwarzaniu informacji i sygnałów w systemach cybernetycznych, we współczesnej literaturze naukowej zwanych często agentami, które mogą być zarówno wyodrębnionymi modułami oprogramowania, procesami w systemach komputerowych (w obu przypadkach agenci tego typu są określane angielskim terminem *disembodied agents*), jak też systemami fizycznymi (ang. *embodied agents*). System cybernetyczny jest dynamiczną, wyodrębnioną jednostką, która działa w pewnym środowisku. Za pomocą wyspecjalizowanych modułów zwanych wejściami (ang. *inputs*) agent odbiera ze środowiska sygnały które niosą pewne informacje. W przypadku agentów fizycznych, w szczególności biologicznych, pobiera też energię i, w niektórych przypadkach, materię. Wspomniana analiza przetwarzania informacji i sygnałów w systemach cybernetycznych, a w przypadku agentów fizycznych również energii i, ewentualnie, materii oraz przesyłania tych zasobów pomiędzy poszczególnymi modułami agenta jest prowadzona w kontekście badań nad sterowaniem, czyli takim oddziaływaniem na system, które umożliwi osiągnięcie założonego celu.

U podstaw cybernetyki leży obserwacja, że procesy sterowania mają pewne wspólne cechy, które można badać niezależnie od tego, czy dotyczą maszyn, organizmów żywych, czy systemów społecznych. Ze względu na przedmiot badań wyróżniamy cybernetykę techniczną (Arbib, 1966; Ashby, 1963; Wiener, 1971; Mazur, 1966), biocybernetykę (Arbib, 1966; Ashby, 1963; Wiener, 1971; Tadeusiewicz, 1994) oraz cybernetykę społeczną. Do teoretycznych narzędzi cybernetyki należą: teoria systemów, teoria modelowania, teoria informacji, teoria gier, teoria decyzji. Znaczącymi w cybernetyce gałęziami matematyki są teoria sterowania, teoria układów dynamicznych, rachunek prawdopodobieństwa i logika matematyczna. Cybernetyka ma wspólne obszary badań z matematyką, fizyką, chemią, biologią, medycyną, informatyką, automatyką, elektroniką, mechatroniką, robotyką, bioniką, filozofią, ekonomią, psychologią, kognitywistyką i socjologią. Trudno jednoznacznie wyznaczyć granice cybernetyki, przy czym dodatkową trudność stanowi dynamika tych granic. Przykładowo zagadnienia dotyczące sztucznej inteligencji w latach 50. i 60. XX wieku były zaliczane do cybernetyki.

Dzisiaj umiejscawia się je raczej w obszarze informatyki. Niemniej jednak, w szerokim rozumieniu, informatykę, podobnie jak automatykę, można zaklasyfikować jako naukę należącą do cybernetyki technicznej.

Analiza historyczna pojęcia i obszaru badań

Mimo że za początek cybernetyki, mając na myśli współczesne rozumienie tego słowa, uważa się rok 1948, w którym Norbert Wiener (1894–1964) opublikował książkę *Cybernetics or Control and Communication in the Animal and the Machine* (1948; polskie wydanie: Wiener, 1971), to początków tej nauki można doszukiwać się znacznie dawniej. Jako pierwszy posłużył się omawianym terminem Platon (427–347 p.n.e.) w dialogu *Gorgiasz*. W innym znaczeniu niż literalnie rozumiane sterowanie statkiem użył wyrazu „cybernetyka” André Ampère (1775–1836), określając nim ten zakres polityki, który obejmuje sztukę rządzenia państwem – było to zatem znaczenie bliskie dzisiejszemu rozumieniu cybernetyki społecznej. W zbliżonym znaczeniu po raz pierwszy w piśmie w języku polskim posłużył się terminem „cybernetyka” filozof Bronisław Trentowski (1808–1869) w wydanej w roku 1843 książce *Stosunek filozofii do cybernetyki, czyli sztuki rządzenia narodem*.

Lata 30. XX wieku to okres, w którym rozpoczęły się intensywne badania w obszarach zaliczanych dzisiaj do cybernetyki. Otrzymano wtedy istotne, często przełomowe, wyniki, mimo że do powszechnie uznanego początku cybernetyki (wspomniany rok 1948) pozostawało kilkanaście lat. Wyniki te uzyskano na gruncie badań prowadzonych w dziedzinie biologii, medycyny, psychologii i psychiatrii, techniki oraz matematyki. W tym okresie zdefiniowano również pewne pojęcia, które weszły potem do cybernetyki jako jej pojęcia podstawowe. Tak więc pojęcie homeostazy jako procesu zapewniającego równowagę funkcjonalną w organizmie zdefiniował w roku 1932 amerykański fizjolog, neurolog i psycholog Walter Cannon (1871–1945). Systematyczne badania naukowe nad niektórymi obszarami, należącymi dzisiaj do matematycznych podstaw cybernetyki oraz do cybernetyki technicznej, rozpoczął Alan Turing (1912–1954), rozważając zagadnienia dotyczące obliczalności funkcji matematycznych na gruncie wprowadzonej przez siebie koncepcji maszyny Turinga (Turing, 1936). Aktualnie jest ona powszechnie

przyjętym modelem formalnym efektywnej obliczeniowo procedury. Z badań Turinga płynie wniosek, że istnieją funkcje, których wartości nie da się obliczyć przy pomocy komputera, co jest istotnym ograniczeniem możliwości tych maszyn. Kurt Gödel (1906–1978), austriacki matematyk, od 1940 roku przebywający na stałe w USA, prowadził w latach 30. XX wieku badania nad logicznymi podstawami matematyki i metamatematyką. W roku 1929 w ramach swego doktoratu udowodnił twierdzenie o pełności rachunku predykatów pierwszego rzędu. W roku 1931 dowiódł słuszności dwóch kolejnych fundamentalnych twierdzeń z podstaw matematyki – o niezupełności każdej aksjomatycznej teorii dedukcyjnej, która zawiera aksjomatykę liczb naturalnych, oraz o niesprzeczności takiej teorii. Twierdzenia Gödla mają fundamentalne znaczenie dla problematyki dowodliwości twierdzeń matematycznych przy pomocy komputera (por. Arbib, 1966).

Istotnym postępowaniem w obszarze dziś zaliczanym do cybernetyki technicznej było skonstruowanie pierwszych maszyn liczących, mających co najmniej niektóre funkcjonalności komputerów. Maszyna ABC (Atanasoff-Berry Computer), skonstruowana w USA w 1939 roku przez Johna Atanasoffa i Clifforda Berry'ego, była wyspecjalizowaną maszyną liczącą przeznaczoną do rozwiązywania liniowych równań algebraicznych. Choć nie była programowalna, uważana jest za prototyp komputera, gdyż była pierwszą liczącą maszyną, w której wykorzystano lampy elektronowe do przeprowadzania operacji arytmetycznych oraz kondensatory jako jednostki pamięci. Za pierwszą programowalną maszynę liczącą należy uznać skonstruowany w Niemczech w 1941 roku kalkulator Z3, oparty jednakże nie na lampach elektronowych, ale na elektrycznych przekaźnikach. Program był zapisywany na taśmie perforowanej, dane wprowadzane z klawiatury, a wyniki odczytywane na wyświetlaczu.

Dalszym krokiem w kierunku skonstruowania komputera było zbudowanie w Wielkiej Brytanii w roku 1943 maszyny Colossus, służącej do łamania niemieckich szyfrów. Arytmetyczne i logiczne operacje były wykonywane przy użyciu lamp elektronowych. Colossus był pierwszą elektroniczną, cyfrową, programowalną maszyną liczącą, chociaż jego program nie rezydował w pamięci ani nie był szczytywany z taśmy perforowanej lub kart perforowanych, ale był realizowany poprzez ustawienie licznych przekaźników.

Kolejny postęp w rozwoju komputerów osiągnięto, konstruując w Stanach Zjednoczonych w roku 1945 ENIAC (*Electronic Numerical Interpreter and Calculator*), zbudowany z ok. 18 000 lamp elektronowych różnego typu oraz ponad 1000 ręcznych przełączników. Nie miał pamięci operacyjnej. Pierwotnie programowany przy pomocy przełączania wtyków kablowych, na późniejszym etapie został wyposażony w czytnik kart perforowanych. Maszyna początkowo przeznaczona była do obliczania tablic balistycznych na potrzeby artylerii dalekiego zasięgu, ale w późniejszym okresie wykonywano na niej obliczenia związane z konstrukcją broni jądrowej, projektowaniem tuneli aerodynamicznych, prognozowaniem pogody oraz badaniami nad promieniowaniem kosmicznym.

Istotnym osiągnięciem teoretycznym z zakresu biocybernetyki było opracowanie w roku 1943 przez neuropsychiatrę Warrena McCullocha (1898–1969) i logika Waltera Pittsa (1923–1969) formalnego modelu neuronu oraz sieci złożonych z kilku neuronów wraz z odniesieniami do logiki.

Kolejne przełomowe wydarzenie w zakresie badań teoretycznych miało miejsce za sprawą biologa, Ludwiga von Bertalanffy'ego (1901–1972), który zaprezentował ogólną teorię systemów na seminarium filozoficznym na Uniwersytecie Chicago w 1937 roku. Istotny impuls do rozwoju cybernetyki dała również problematyka obliczeniowa związana z kierowaniem ogniem artylerii przeciwlotniczej. Ze względu na gwałtowny wzrost szybkości samolotów rachunki musiały uwzględniać przyszłe położenie samolotu i w związku z wykonywanymi przez niego manewrami musiały być nieustannie na bieżąco korygowane. Doprowadziło to do opracowania mającej zastosowanie przy sterowaniu idei ujemnego sprzężenia zwrotnego, które potem stało się jednym z podstawowych pojęć cybernetyki.

W pierwszej połowie lat 40. w Wielkiej Brytanii psychiatra Ross Ashby (1903–1972) badał analogie między procesami adaptacyjnymi zwierząt i ludzi a stanami równowagi w układach mechanicznych, kontynuując prace Cannona nad homeostazą. Prace te zostały zwieńczone skonstruowaniem przez Ashby'ego elektromechanicznego homeostatu w roku 1948. Od roku 1946 z inicjatywy McCullocha zaczęły odbywać się w Nowym Jorku coroczne interdyscyplinarne konferencje poświęcone problematyce „Sterowania i łączności w maszynach i zwierzętach”,

a w 1947 roku Wiener nazwał nowo powstałą naukę zwaną cybernetyką. W tym samym czasie, w 1948 roku, Claude Shannon (1916–2001) stworzył matematyczną teorię komunikacji (Shannon, 1948), rok później rozszerzoną do teorii informacji i kodowania. Jak już wspomniano, rok opublikowania słynnej książki Wienera wraz z zaproponowaniem teorii komunikacji i informacji uważa się powszechnie za moment powstania nowej nauki – cybernetyki. U podstaw cybernetyki legły również prace Johna von Neumanna (1903–1957), węgierskiego matematyka, fizyka, informatyka i cybernetyka pracującego w Stanach Zjednoczonych. W cybernetyce stworzył on teorię automatów (von Neumann, 1951), prowadząc badania nad automatami samoreplikującymi się oraz definiując automat komórkowy. Od tego momentu cybernetyka intensywnie się rozwijała, przy czym można zaobserwować kilka nurtów badań.

Pierwszym nurtem badań, ściśle związanym z rozwojem elektroniki, był rozwój techniki obliczeniowej. Zgodnie z powszechnie przyjętą klasyfikacją komputerów ze względu na użytą technikę jego konstrukcji wyróżniamy komputery zerowej generacji – oparte na technice przekąźnikowej; pierwszej generacji – komputery oparte na lampach elektronowych; drugiej generacji – komputery tranzystorowe; trzeciej generacji – komputery na układach scalonych małej skali integracji; czwartej generacji – komputery na układach scalonych średniej i dużej skali integracji; piątej generacji – komputery oparte na innej technice niż krzemowa, np. komputery kwantowe, optyczne, chemiczne, biochemiczne. Na dzień dzisiejszy komputery piątej generacji są w początkowej fazie eksperymentalnej.

Po skonstruowaniu ENIAC-a z jednej strony komputery lampowe były systematycznie udoskonalane i wchodziły w użycie na uczelniach. Z drugiej strony sukcesy w rozwoju techniki półprzewodnikowej, w szczególności budowa diody półprzewodnikowej i tranzystora, zapoczątkowały prace nad komputerami drugiej generacji. Pierwszy tranzystor – tranzystor punktowy (ostrzowy) – został skonstruowany w 1947 roku w laboratorium amerykańskiej firmy Bell Telephone Laboratories. Tranzystor punktowy szybko został wyparty przez skonstruowany w 1950 roku tranzystor warstwowy (tranzystor złączowy). Pierwszym komputerem częściowo drugiej generacji był zbudowany w 1953 roku na Uniwersytecie w Manchesterze komputer, do konstrukcji którego użyto zarówno tranzystorów, jak również lamp elektronowych. W Bell

Telephone Laboratories w roku 1954 na potrzeby bombowców B-52 został skonstruowany komputer pokładowy, również oparty częściowo na technice półprzewodnikowej, a częściowo na lampowej.

Pierwszy w pełni półprzewodnikowy komputer drugiej generacji został zbudowany w Wielkiej Brytanii w Zakładzie Badań Energii Atomowej w 1955 roku. Dysponował on jednak mniejszą mocą obliczeniową niż produkowane wówczas komputery lampowe. Dominacja komputerów drugiej generacji rozpoczęła się w roku 1959, gdy skonstruowano komputer IBM 7090, którego moc obliczeniowa była większa niż ówczesnych komputerów lampowych. Komputer ten był wykorzystywany jeszcze w latach 80. XX wieku.

Pierwszym komputerem trzeciej generacji był IBM 360, skonstruowany w 1965 roku. Osiągnięta przy pomocy układów scalonych dalsza – w stosunku do komputerów tranzystorowych – miniaturyzacja i zmniejszenie poboru mocy pozwoliły na skonstruowanie minikomputerów, które z kolei stały się katalizatorem opracowania nowego typu systemów operacyjnych.

Komputery czwartej generacji zawierają mikroprocesor, czyli pojedynczy układ scalony o wielkim stopniu integracji, zdolny do wykonywania operacji wyspecyfikowanych w postaci ciągu instrukcji. Mikroprocesor jest centralną jednostką komputera (CPU – Central Processing Unit). Pierwszy mikroprocesor został zbudowany w 1971 roku. Technologia charakteryzująca czwartą generację komputerów pozwoliła na budowę mikrokomputerów i, w konsekwencji, komputerów osobistych. Pojawiły się też komputery wielordzeniowe, wprowadzone na rynek w 2001 roku. Równocześnie opracowanie wyświetlaczy ciekłokrystalicznych oraz tranzystorów cienkowarstwowych, mogących sterować takimi wyświetlaczami, pozwoliło na skonstruowanie laptopów. Wzrost szybkości obliczeń oraz rozwój techniki telekomunikacyjnej umożliwiły, z kolei, powstanie sieci komputerowych i, w konsekwencji, Internetu.

Drugi nurt badań cybernetycznych związany był ze sztuczną inteligencją. W latach 50., gdy moc obliczeniowa komputerów była niewielka w porównaniu z potrzebami systemów sztucznej inteligencji, badania z konieczności koncentrowały się na aspekcie teoretycznym, a prototypowe implementacje charakteryzowały się prostotą. Pierwszym historycznie nurtem badań w tym obszarze były próby tworzenia formalnych modeli wnioskowania przeprowadzanego przez człowieka. W roku 1955

zaimplementowano system Logic Theorist, który służył do dowodzenia twierdzeń matematycznych. Wykazał się on skutecznością, dowodząc kilkudziesięciu znanych twierdzeń matematycznych. Pod koniec lat 50. XX wieku rozpoczęto również prace nad abstrakcyjnymi metodami wnioskowania opartymi na logice formalnej, a nie na modelowaniu myślenia.

W roku 1957 miało miejsce również istotne osiągnięcie w nurcie badań nad sieciami neuronowymi – zaimplementowano pierwszy elektromechaniczny system symulujący działanie niewielkiej sieci neuronowej – tzw. perceptron, skonstruowany przez amerykańskiego psychologa Franka Rosenblatta (1928–1971). Była to pierwsza sztuczna sieć neuronowa. Neurony w perceptronie były skonstruowane zgodnie z modelem McCullocha i Pittsa, a sama sieć miała strukturę warstwową. Oznacza to, że zbiór wszystkich neuronów w sieci był podzielony na podzbiory, odpowiadające poszczególnym warstwom. Neuron w danej warstwie był połączony ze wszystkimi neuronami kolejnej warstwy i z żadnymi innymi. Połączenie neuronów oznacza, że sygnał wyjściowy jednego neuronu jest równocześnie sygnałem wejściowym drugiego neuronu. Perceptron modelował działanie systemu wzrokowego, w którym połączenia neuronalne tworzą strukturę warstwową. Badania Rosenblatta koncentrowały się na możliwościach nauki tego typu sieci, czyli algorytmu ustawienia wag poszczególnych neuronów, które to wagi są parametrami systemu. Prowadzone również były badania nad możliwością realizacji operacji logicznych przy pomocy pojedynczego neuronu, jak również sieci warstwowych. Lata 80. i 90. XX wieku to okres intensywnego rozwoju badań nad sztucznymi sieciami neuronowymi. Zaproponowano wtedy nowe rodzaje sieci oraz wiele algorytmów ich nauki. Otrzymano też cenne wyniki w zakresie zarówno analizy nauki sieci neuronowych, aproksymacyjnych własności sieci, jak również dynamiki rekurencyjnych sieci neuronowych oraz możliwości zastosowań. Opisano także nowe modele neuronu.

Wiek XXI owocuje kolejnymi rodzajami sieci neuronowych, m.in. sieciami głębokimi oraz grafowymi. Pod koniec lat 60. i w latach 70. XX wieku tematyka sztucznej inteligencji wzbogaciła się o badania nad reprezentacją wiedzy. Zaproponowano kilka typów sieci semantycznych, w których wiedza była modelowana w postaci grafu. Jego wierzchołki odpowiadały obiektom, a krawędzie – relacjom między obiektami.

Wysunięto propozycję opisanego systemów skryptowych jako formalnego modelu języka naturalnego opisującego strukturę świata i jego semantykę. Z badań nad językami naturalnymi wyłoniły się na gruncie cybernetyki syntaktyczne metody rozpoznawania obiektów i analizy sceny, na której operuje agent. Polegają one na formalnym opisie prawideł konstrukcji obiektów. Kompletny zbiór takich prawideł tworzy opisującą gramatykę, generującą wszystkie wzorce sceny. Od strony teoretycznej gramatyka jest formalnym opisem języka, do którego zdań należą analizowane obiekty. Analiza ta jest dokonywana przy pomocy automatu formalnego, który rozstrzyga, czy dany obiekt należy do języka generowanego przez zdefiniowaną gramatykę. Metody syntaktyczne są wykorzystywane m.in. w systemach wizyjnych robotów autonomicznych oraz w analizie obrazów medycznych.

Bardzo ciekawym nurtem badań inspirowanym biologią jest modelowanie procesów ewolucyjnych, rozpoczęte pod koniec lat 50. Przełomowe były wyniki Johna Hollanda (1929–2015), amerykańskiego psychologa, informatyka i matematyka z Uniwersytetu w Michigan, który jako pierwszy zaimplementował algorytm genetyczny. Modele takie, jak większość modeli inspirowanych biologicznie, mają istotną wartość, funkcjonując z jednej strony jako narzędzie teoretyczne służące do analizy danego zjawiska biologicznego. Z drugiej strony służą jako punkt wyjścia do stworzenia sztucznych systemów, mających pewne funkcjonalne własności biologicznych pierwowzorów i ten aspekt badań jest ściśle cybernetyczny. Trzecim ich aspektem jest fakt, że niezależnie od biologicznych korzeni stają się często skuteczną podstawą do opracowania algorytmów rozwiązujących szeroką gamę problemów informatycznych. Do tego typu modeli należą również sztuczne systemy immunologiczne wprowadzone w roku 1986 przez amerykańskich fizyków J. Doyne'a Farmera (ur. 1952), Normana Packarda (ur. 1954) oraz informatyka i biologa Alana Perelsona. Od początku wprowadzenia modelu systemu immunologicznego badano możliwości jego wykorzystania jako systemu sztucznej inteligencji. Obecnie sztuczne systemy immunologiczne są wykorzystywane jako klasyfikatory wzorców, modele pamięci asocjacyjnej oraz detektory anomalii. Należy podkreślić, że badania związane ze sztuczną inteligencją wykraczają poza obszar cybernetyki i mają istotne aspekty dotyczące dziedzin filozofii i psychologii (Flasiński, 2011).

Kolejnym zagadnieniem związanym z cybernetycznym modelowaniem procesów i struktur biologicznych są obliczenia membranowe (systemy membranowe, P-systemy), zaproponowane w roku 1998 przez rumuńskiego informatyka Gheorghe'a Păuna (ur. 1950). Podstawą systemu obliczeniowego jest zbiór membran tworzących częściowy porządek generowany przez relację zawierania się jednej membrany w innej. Obliczenia są wykonywane równolegle wewnątrz poszczególnych membran, a po zakończeniu etapu obliczeń wyniki są przesyłane do innej membrany. Struktura i dynamika procesów obliczeniowych w systemach membranowych jest zorganizowana na wzór procesów biochemicznych i biofizycznych w układach biologicznych. Badania cybernetyczne, inspirowane biologicznie, ale niezwiązane ze sztuczną inteligencją, dotyczą scenariuszy powstania życia oraz badań nad minimalnymi systemami spełniającymi definicję układu żywego. Do najbardziej znanych współczesnych nurtów badawczych tego typu należą badania związane z *chemotone* – teoretycznym konstruktem modelującym minimalną sieć procesów chemicznych przejawiającą cechy układu żywego. Chemoton został zaproponowany w roku 1974 przez węgierskiego biologa i chemika Tibora Gántiego (1933–2009) (Gánti, 1975).

Trzeci główny nurt badań cybernetycznych związany jest z robotyką, zwłaszcza biologicznie inspirowanymi robotami autonomicznymi. Już we wczesnej fazie rozwoju cybernetyki próbowano konstruować roboty, które imitowały odruchy biologiczne. Pierwszymi tego typu konstrukcjami były *żółwie cybernetyczne*, skonstruowane przez brytyjskiego neurofizjologa i cybernetyka Williama Graya Waltera (1910–1977) w roku 1949. Były to pierwsze roboty, które wykazywały zachowania autonomiczne. W najprostszej wersji wyposażone były w czujnik dotyku i czujnik światła. Zaprogramowane były tak, żeby wykazywać reakcje na bodźce świetlne, dążąc do źródła światła (znana z biologii fototaksja dodatnia). W pierwszej wersji zbudowane zostały dwa egzemplarze – Elmer i Elsie. Pierwsze eksperymenty polegały m.in. na zbadaniu wpływu wzmocnienia bodźców na działanie robota. Okazało się, że zbytne wzmocnienie bodźców wpływa niekorzystnie na szybkość odnalezienia źródła światła.

Kolejnymi etapami rozwoju robotyki były zapoczątkowane w latach 60. XX wieku badania nad systemami wizyjnymi robotów oraz zastosowanie

robotów przemysłowych w fabrykach jako element linii produkcyjnych (lata 70.). Począwszy od lat 70. roboty, jak dotąd wyłącznie zdalnie sterowane, używane były do badania powierzchni Księżyca (radzieckie łunochody) oraz – począwszy od lat 90. – Marsa. Dziś konstruuje się roboty o bardzo rozbudowanych zdolnościach motorycznych, zarówno człekokształtne, jak też wzorowane na zwierzętach. Przoduje na tym polu amerykańska firma Boston Dynamics.

Na szczególną uwagę zasługuje polska szkoła cybernetyki. Do połowy lat 50. XX wieku cybernetyka w ZSRR i, co za tym idzie, również w Polsce, uważana była za reakcyjną pseudonaukę i w związku z tym wszelkie badania z nią związane musiały być prowadzone bez używania terminu „cybernetyka” i z nikłą możliwością publikacji wyników. Od roku 1956 uprawianie cybernetyki było już w Polsce możliwe i mimo obiektywnie niekorzystnych warunków do rozwoju, związanych z ograniczoną możliwością wymiany informacji naukowej w skali międzynarodowej w czasach komunizmu – zwłaszcza w latach 50. i 60. XX wieku – oraz z bardzo ograniczoną możliwością finansowania badań naukowych w Polsce komunistycznej, rozwijała się dynamicznie. Do pierwszego pokolenia polskich cybernetyków należał Henryk Greniewski (1903–1972). Był logikiem i matematykiem specjalizującym się w matematyce finansowej. Od roku 1948 zaangażowany w prace nad pierwszymi polskimi komputerami, stał się współtwórcą (wraz z matematykami Kazimierzem Kuratowskim i Andrzejem Mostowskim) Grupy Aparatów Matematycznych przy Państwowym Instytucie Matematycznym w Warszawie. Wprowadził pojęcie „systemu względnie odosobnionego”, odpowiadającego współcześnie używanemu terminowi „system cybernetyczny” lub „agent”. Zajmował się modelowaniem odruchów warunkowych, badał kody i języki, analizował teorię informacji i teorię sieci. Opracowywał modele zachowania zwierząt – tzw. zwierzęta syntetyczne. W roku 1962 z jego inicjatywy powstało Polskie Towarzystwo Cybernetyczne. Od 1946 roku pracował na Uniwersytecie Jagiellońskim Mieczysław Choynowski (1909–2001) psycholog i filozof twórca polskiej psychometrii. Zajmował się między innymi modelowaniem działania układu nerwowego, stosując metody cybernetyki, przede wszystkim koncepcję ujemnego sprzężenia zwrotnego i homeostazy. Pod jego kierunkiem w Konwersatorium Naukoznawczym, działającym przy Uniwersytecie Jagiellońskim w latach 1945–1950, pracował w latach

1948–1950 Stanisław Lem (1921–2005), co miało istotne znaczenie zarówno w jego późniejszej eseistyce naukowej dotyczącej cybernetyki, jak i w twórczości literackiej z gatunku *science fiction*.

Choynowski współpracował z Antonim Kępińskim (1918–1972), twórcą oryginalnej koncepcji metabolizmu informacyjnego, według której w układach biologicznych przetwarzanie informacji jest istotną częścią metabolizmu. W związku z tym analiza i modelowanie klasycznie rozumianego metabolizmu, czyli całokształtu reakcji chemicznych i związanych z nimi przemian energetycznych w komórkach, muszą być uzupełnione o komponentę informacyjną. Kępiński zajmował się również analizą reprezentacji poznawczych – tzw. ramy poznawcze oraz dynamiką reakcji psychologicznych. Należy podkreślić, że nie uważał on swojej teorii za podejście cybernetyczne. Źródłem jego rezerwy w tym względzie był prawdopodobnie fakt, że z jednej strony cybernetykę uważał za naukę techniczną, a z drugiej strony był przeciwnikiem wszelkiego redukcjonizmu w zastosowaniu do człowieka, umieszczając swoje poglądy w nurcie personalizmu.

Współtwórcą cybernetyki ekonomicznej był Oskar Lange (1904–1965), jedyny reprezentant pierwszego pokolenia polskich cybernetyków, którego prace były szeroko znane w świecie zachodnim. Analizował problem centralizacji i decentralizacji w zarządzaniu, zajmował się procesami regulacji i samoregulacji w układach ekonomicznych oraz związkami cybernetyki i filozofii. Wspomniany już Stanisław Lem w swojej eseistyce naukowej poświęcał uwagę granicom możliwości modelowania układów biologicznych, w tym aspektów psychologicznych, rzeczywistości wirtualnej, którą nazywał „fantomatyką” (Lem, 1974), własnościom umysłu i jego ograniczeniom wynikającym z uwarunkowań biologicznych, jak również możliwości przekroczenia tych ograniczeń w systemach sztucznej inteligencji. Zajmował się również socjologicznymi aspektami rozwoju cybernetyki. Unikalna do dnia dzisiejszego jest jego cybernetyczna analiza dzieła literackiego. Oryginalne analizy i koncepcje z obszaru cybernetyki przedstawiał również w swoich powieściach. Mimo że poza epizodem pracy w Konwersatorium Naukoznawczym Lem nie był naukowcem, to był i jest cytowany w pracach ściśle naukowych nie tylko filozoficznych, ale również cybernetycznych, przy czym cytowane są nie tylko jego eseje naukowe, ale także powieści, ze względu na zawarty w nich ładunek cybernetycznych idei i analiz.

Marian Mazur (1909–1983) jest twórcą oryginalnej teorii układów autonomicznych oraz teorii informacji jakościowej (Mazur, 1966). Swoją koncepcję układu autonomicznego zastosował do analizy struktury i dynamiki ludzkiej psychiki. Zajmował się również cybernetyką ekonomiczną i społeczną, którą to tematykę kontynuował Józef Kossecki (1936–2005), prowadząc badania m.in. w zakresie cybernetyki kultury.

Neurofizjolog Jan Trąbka (1931–2012), kierownik Zakładu Biocybernetyki Collegium Medicum UJ w latach 1991–2002, zajmował się neurocybernetyką. W latach 1959–1960, jako stypendysta Fundacji Rockefellera, pracował na Uniwersytecie Harvarda, współpracując z Norbertem Wienerem. Prowadził badania nad mózgiem, między innymi analizując metodami cybernetycznymi zjawisko świadomości. W jego badaniach bardzo istotny był filozoficzny aspekt cybernetyki, którą rozumiał znacznie szerzej niż naukę o sterowaniu i informacji, traktując cybernetykę jako fenomen światopoglądowy.

Ryszard Tadeusiewicz (ur. 1947), biocybernetyk i informatyk, pracownik AGH i jej rektor w latach 1998–2005, prowadzi intensywne badania nad systemami sztucznej inteligencji, w szczególności nad sztucznymi sieciami neuronowymi. Tematem jego prac były też systemy wizyjne w kontekście robotyki przemysłowej. Cenne wyniki uzyskał na gruncie modelowania układów biologicznych, przede wszystkim w neurocybernetyce (Tadeusiewicz, 1994), ale również w modelowaniu zachowań roju owadów. Wniósł istotny wkład w dziedzinę rozpoznawania obrazów (*pattern recognition*) oraz analizy i rozumienia sceny, zwłaszcza w kontekście robotyki oraz analizy obrazów medycznych. Analizował też model układu autonomicznego, zaproponowany przez Mariana Mazura, badając możliwości implementacji takiego układu i jego zachowanie w zmiennym środowisku.

Zdzisław Pawlak (1926–2006) jest twórcą teorii zbiorów przybliżonych, formalnego modelu maszyny liczącej zwanej maszyną Pawlaka oraz modelu matematycznego kodów genetycznych DNA, opracował też oryginalne matematyczne ramy teorii konfliktów. W roku 1950 zbudował komputer GAM-1 – pierwszy komputer powstały w Grupie Aparatów Matematycznych Państwowego Instytutu Matematycznego. Komputer oparty był na technologii przekaźnikowej.

Aktualnie polscy naukowcy biorą czynny udział w badaniach nad sztuczną inteligencją, biocybernetyką oraz teorią informacji. W systemach sztucznej inteligencji polskie środowisko naukowe szczególnie

aktywne jest w dziedzinie teoretycznych podstaw, sztucznych sieci neuronowych oraz metod syntaktycznych.

Problemowe ujęcie pojęcia i obszaru badań

Cybernetyka bada działanie pewnej klasy systemów egzystujących w swoim środowisku, posługując się modelami teoretycznymi oraz konstruując sztuczne systemy. Na poziomie podstawowych problemów cybernetycznych badane są na przykład schematy bodziec-reakcja danego systemu cybernetycznego. Jedną z metodologii stosowanych w badaniach tego typu bazuje na pojęciu czarnej skrzynki. Metodologia ta została zapożyczona z elektrotechniki, gdzie została zaproponowana i zastosowana po raz pierwszy w 1941 roku. Na potrzeby cybernetyki została zaadaptowana przez Ashby'ego w 1956 roku (Ashby, 1963). Czarna skrzynka jest systemem, którego struktura wewnętrzna jest badaczowi nieznaną, podobnie jak cele i algorytmy, które determinują jego działanie. Badacz może jednak oddziaływać na system poprzez jego wejścia i obserwować działanie systemu, w szczególności jego sygnały wyjściowe. Po uzyskaniu dostatecznie dużej próbki reakcji systemu na różne bodźce eksperymentator może próbować znaleźć reguły transformacji bodziec-reakcja dla danego systemu lub przynajmniej ocenić, czy schemat bodziec-reakcja jest deterministyczny. Cybernetyka dostarcza ogólnych pojęć i formalnych technik do badania ogólnych własności systemów cybernetycznych, ich dynamiki i struktury niezależnie od nośnika, na którym systemy te są osadzone. A zatem niezależnie od tego, czy jest to system techniczny, biologiczny czy socjalny, możemy badać jego poziom stabilności, strukturę przepływu sygnałów i energii, sposób, w jaki sygnały konstytuują i kodują w systemie informację, w szczególności tę jej część, która jest związana ze sterowaniem, czy sposób oddziaływania systemu na swoje otoczenie, przede wszystkim w kontekście realizacji celów. Możemy więc analizować strukturę homeostatu, schemat jego sprzężenia z pozostałymi modułami systemu, sposób realizacji homeostazy, czyli strukturę i siłę sprzężeń zwrotnych zapewniających homeostazę. Tego typu metodyka pozwoliła na zaproponowanie uniwersalnej struktury systemu autonomicznego oraz jego szczegółową analizę (Mazur, 1966).

We współczesnych badaniach naukowych cybernetyka odgrywa istotną rolę. Przede wszystkim jest powszechnie stosowaną metodologią w biologii. Powstała nowa, dynamicznie rozwijająca się gałąź biologii – biologia systemów, która bada procesy biologiczne i dynamikę biologicznych struktur metodami cybernetyki. Istotną tematyką badawczą jest przepływ i kodowanie informacji w układach biologicznych. Powszechnie przyjętym paradygmatem jest obecnie stwierdzenie, że informacja i jej przetwarzanie leżą u podstaw życia. Stwierdzenia, że „życie = materia + informacja” lub „nic w biologii nie ma sensu, dopóki nie weźmiemy pod uwagę przepływu informacji” mają w biologii współczesnej rangę stwierdzeń oczywistych. Cybernetyczna metodologia odgrywa istotną rolę również w fundamentalnych problemach biologii. Kontynuowane są wspomniane wyżej badania nad możliwymi scenariuszami początków życia oraz nad istotą życia jako takiego. W ramach tej ostatniej tematyki proponowane są cybernetyczne definicje życia oraz cybernetyczna analiza zjawiska życia (Korzeniewski, 2001), których podstawowym założeniem jest, że życie w swojej istocie wychodzi poza paradygmat biologiczny i, w związku z tym, chcąc zbadać jego istotę, musimy wyjść poza metody czysto biologiczne. W tym kontekście prowadzone są również badania nad informacją w biologii i chemii (Bielecki & Schmittel, 2022). Z informacją w biologii łączy się kolejna współcześnie powstała gałąź cybernetyki – cybernetyka molekularna (Schmittel, 2019). Jest to obszar chemii supramolekularnej. Tematyka badawcza cybernetyki molekularnej obejmuje zagadnienia przełączników chemicznych, chemicznych bramek logicznych i maszyn molekularnych. Rozwój Internetu stymuluje badania cybernetyczne związane z informatyką. Prowadzone są badania nad rozwojem inteligentnych asystentów wspomagających pracę z Internetem, realizujących na przykład inteligentne wyszukiwanie w sieci. Tworzone i rozwijane są inteligentne systemy dialogiczne, na przykład ChatGPT. Intensywnie usprawniane są systemy wieloagentowe, w kontekście których wyłoniła się tematyka badawcza związana z tzw. agentami emocjonalnymi. Ma ona dwa aspekty. Pierwszy, obecnie bardziej popularny, wiąże się z wykrywaniem ludzkich emocji przez systemy sztucznej inteligencji współpracujące bezpośrednio z człowiekiem, aby móc aktualną formę działania najlepiej dostosować do stanu emocjonalnego użytkownika. Drugi aspekt to modelowanie emocji w systemach sztucznej inteligencji, na przykład w robotach autonomicznych,

aby zaimplementowane algorytmy modelujące emocje zapewniały analogiczną funkcjonalność systemu jak emocje w układach biologicznych, pozwalając na przykład robotowi autonomicznemu wyposażonemu w symulator lęku unikać sytuacji potencjalnie zagrażających jego egzystencji. Robotyka autonomiczna wymaga coraz bardziej złożonej sensoryki. Integracja bodźców z sensorów różnego typu i zbudowanie z otrzymanych informacji integralnej wizji środowiska pozwalającej na jego „rozumienie” przez sztuczny system cybernetyczny należy do aktualnych wyzwań w dziedzinie cybernetyki.

Refleksja systematyczna i wnioski

Cybernetyka to niezwykle szeroki obszar nauki, który wyewoluował w automatykę, informatykę i robotykę, uważane dziś za odrębne dyscypliny naukowe. Z drugiej strony niekiedy są one zaliczane do cybernetyki, a powszechnie przyjmuje się, że mają z nią szerokie obszary wspólne. Fakt nie do końca określonych granic cybernetyki manifestuje się w szczególności w tym, że cybernetyka nie funkcjonuje w Polsce w oficjalnym wykazie dziedzin i dyscyplin naukowych. W latach 2005–2018 w dziedzinie nauk technicznych wyspecyfikowana była jedynie „biocybernetyka i inżynieria biomedyczna”, zastąpiona potem przez „inżynierię biomedyczną”. Z drugiej jednak strony istnieją prestiżowe czasopisma naukowe poświęcone cybernetyce zarówno w całościowym aspekcie, jak i jej poszczególnym gałęziom, przede wszystkim biocybernetyce. Na wielu uczelniach istnieją instytuty cybernetyki.

Podsumowując, cybernetyka intensywnie się rozwija, posiadając wspólne obszary badawcze z wieloma innymi naukami. Coraz powszechniej przenikająca do codziennego życia robotyka, tzw. inteligentne środowiska – inteligentne samochody, inteligentne domy, inteligentne sieci energetyczne, Internet oferujący coraz więcej funkcji, stymulują rozwój cybernetyki zarówno w aspekcie teoretycznym, jak i aplikacyjnym. Powszechność systemów cybernetycznych ma również aspekt kulturowy. Takie pojęcia jak „cyberbezpieczeństwo”, „cybersieć” czy „cyberseks” weszły na stałe do potocznego języka. Uzależnienie od gier komputerowych lub cyberseksu jest coraz większym problemem społecznym i medycznym. Widać więc, że cybernetyka ma też swoją ciemną stronę.

BIBLIOGRAFIA

- Arbib, M.A. (1966) *Mózg, maszyna, matematyka*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Ashby, W.R. (1963). *Wstęp do cybernetyki*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Bielecki, A., & Schmittel, M. (2022). The information encoded in structures: Theory and applications to molecular cybernetics. *Foundations of Science*, 27(2), 1327–1345.
- Flasiński, M. (2011). *Wstęp do sztucznej inteligencji*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Gánti, T. (1975). Organisation of chemical reactions into dividing and metabolizing units: the chemotons. *BioSystems*, 7(1), 189–195.
- Korzeniewski, B. (2001). Cybernetic formulation of the definition of life. *Journal of Theoretical Biology*, 209(3), 275–286.
- Lange, O. (1965). *Wstęp do cybernetyki ekonomicznej*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Lem, S. (1974). *Summa technologiae*. Kraków: Wydawnictwo Literackie.
- Mazur, M. (1966). *Cybernetyczna teoria układów samodzielnych*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Shannon, C.E. (1948). The Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3), 379–423.
- Schmittel, M. (2019). Dynamic functional molecular systems: From supramolecular structures to multi-component machinery and to molecular cybernetics. *Israel Journal of Chemistry*, 59(3), 197–208.
- Tadeusiewicz, R. (1994). *Problemy biocybernetyki*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Turing, A.M. (1936). On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(1), 230–265.
- von Neumann, J. (1951). The general and logical theory of automata. W: *Cerebral Mechanisms and Behavior: The Hixton Symposium*. New York: John Wiley and Sons.
- Wiener, N. (1971). *Cybernetyka, czyli sterowanie i komunikacja w zwierzęciu i maszynie*. Warszawa: Państwowe Wydawnictwo Naukowe.

Andrzej Bielecki
AGH Akademia Górniczo-Hutnicza
<https://orcid.org/0000-0002-0192-3785>
<https://doi.org/10.35765/slowniki.418>

Sztuczne sieci neuronowe

Streszczenie

DEFINICJA POJĘCIA: Sztuczna sieć neuronowa jest systemem, którego struktura i działanie są wzorowane na układzie nerwowym. Podstawowa jednostka funkcjonalna sztucznej sieci neuronowej to sztuczny neuron, będący modelem neuronu biologicznego. Sieć jest zbudowana w ten sposób, że wyjścia jednych neuronów są połączone z wejściami innych neuronów. W zależności od schematu połączeń otrzymujemy różne struktury sieci.

ANALIZA HISTORYCZNA POJĘCIA: U podstaw modelowania układów neuronalnych leżą wyniki badań neurofizjologów uzyskiwane od drugiej połowy XIX wieku. Założenie, że neuron jest podstawową jednostką funkcjonalną układu nerwowego, pozwoliło w 1943 roku opracować pierwszy teoretyczny model neuronu oraz zaimplementować pierwszą sztuczna sieć neuronową w roku 1957, która była siecią warstwową. Opracowanie teoretycznych podstaw nauki takich sieci, zaproponowanie nowych rodzajów zarówno samych sieci, jak i metod ich nauki w latach 80. XX wieku oraz postępy w matematycznej analizie ich działania spowodowały ich szybki rozwój pod koniec XX wieku oraz w wieku XXI.

UJĘCIE PROBLEMOWE POJĘCIA: Sztuczna sieć neuronowa może być zaimplementowana zarówno jako oprogramowanie, jak również jako układ elektroniczny, analogowy lub cyfrowy. Istotnym zagadnieniem jest nauka sieci, czyli ustawienie jej parametrów. Na ogół nauka jest realizowana przy pomocy algorytmu iteracyjnego, który wykorzystuje przygotowane przez programistę przykłady sygnałów wejściowych. Sztuczne sieci neuronowe są wykorzystywane do modelowania różnych aspektów układu nerwowego i procesów psychicznych.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI:

Niezależnie od biologicznych inspiracji sztuczne sieci neuronowe są używane jako pewien typ systemu sztucznej inteligencji, który skutecznie rozwiązuje szeroką gamę zadań. Należą do nich problemy optymalizacji, klasteryzacji, sterowania, rozpoznawania wzorców, diagnostyki i predykcji.

Słowa kluczowe: sztuczna sieć neuronowa, model neuronu,
nauka sieci neuronowej

Definicja pojęcia i określenie obszaru badań

Sztuczna sieć neuronowa (ang. *artificial neural network*) jest oprogramowaniem lub systemem elektronicznym, którego budowa i działanie wzorowane są na budowie i działaniu układu nerwowego. Podstawowym założeniem, będącym punktem wyjścia modelowania układów neuronalnych i – co za tym idzie – podstawą konstrukcji sztucznych sieci neuronowych, jest tzw. teoria neuronalna, zgodnie z którą neurony są podstawowymi jednostkami funkcjonalnymi układu nerwowego. W związku z tym pierwszym krokiem implementacji, zarówno w postaci oprogramowania, jak i w postaci układu elektronicznego, jest stworzenie formalnego modelu neuronu. Pierwszy taki model został opracowany w roku 1943 przez dwóch amerykańskich naukowców – neurofizjologa i cybernetyka Warrena McCullocha (1899–1969) oraz logika Waltera Pittsa (1923–1969), którzy zaproponowali zarówno model pojedynczego neuronu, jak również model sztucznej sieci neuronowej zbudowanej z takich neuronów. Badali także możliwości realizowania przez takie sieci obliczeń logicznych klasycznego rachunku zdań (McCulloch & Pitts, 1943).

Model neuronu McCullocha-Pittsa był w oryginalnej wersji jednostką progową z kilkoma ważonymi wejściami i jednym wyjściem. W momencie gdy całkowite wzbudzenie neuronu przekroczyło zadaną wartość progową, na wyjściu neuronu pojawiał się znormalizowany sygnał. Całkowite wzbudzenie neuronu oblicza się jako sumę iloczynów sygnałów wejściowych na poszczególnych wejściach neuronu przez wartość wagi tego wejścia. Wagi są traktowane jako parametry sztucznego neuronu, którym należy nadać taką wartość, aby zaimplementowany układ, czyli sztuczny neuron lub sieć złożona z takich neuronów działała, jako układ obliczeniowy, zgodnie z oczekiwaniami projektanta, rozwiązując postawione zadanie.

Sztuczna sieć neuronowa to układ złożony ze sztucznych neuronów połączonych ze sobą w ten sposób, że sygnał wyjściowy danego neuronu jest podawany na wejście innych neuronów. Pierwsza implementacja sztucznej sieci neuronowej została wykonana w roku 1957 przez amerykańskiego psychologa, Franka Rosenblatta (1928–1971), zajmującego się sztuczną inteligencją (Rosenblatt, 1958). Była to sieć warstwowa, czyli taka sztuczna sieć neuronowa, w której neurony tworzą

warstwy, przy czym neurony z danej warstwy połączone są tylko z neuronami kolejnej warstwy. Pierwowzorem perceptronu, bo tak Rosenblatt nazwał zaimplementowaną przez siebie sztuczną sieć neuronową, był układ wzrokowy, a konkretnie budowa siatkówki, w której neurony tworzą warstwy.

Począwszy od lat 70. XX wieku obszar badań naukowych związany ze sztucznymi sieciami neuronowymi intensywnie się rozwija. Jednym z najważniejszych zagadnień badawczych jest problem nauki sieci neuronowych, zwany też trenowaniem sieci, czyli opracowanie algorytmu dopasowywania wag poszczególnych neuronów w sieci. W niektórych algorytmach nauki zachodzi również dodatkowo usuwanie poszczególnych neuronów lub dodawanie neuronów, co oznacza, że sieć w trakcie nauki zmienia swoją strukturę (tzw. ontogeniczne sieci neuronowe). Odpowiada to procesom biologicznym związanym z neuroplastycznością mózgu, która polega przede wszystkim na rozbudowie siatki połączeń między neuronami, ale też na zmianie liczby neuronów. Kolejnym istotnym tematem badawczym jest opracowywanie nowych rodzajów sztucznych sieci neuronowych, opartych na różnych modelach neuronów i mających także różnorakie struktury. Intensywnie badane są matematyczne podstawy sztucznych sieci neuronowych oraz możliwości ich zastosowań zarówno jako modeli różnych neuronalnych struktur biologicznych, jak również – abstrahując od ich biologicznych korzeni – zastosowania sieci neuronowych jako systemów sztucznej inteligencji używanych do rozwiązywania szerokiej gamy problemów.

Analiza historyczna pojęcia i obszaru badań

U podstaw implementacji sztucznych sieci neuronowych leżą badania nad układem nerwowym. Tkanka nerwowa składa się z trzech komponentów – neuronalnego, glistowego i naczyniowego. Komponent neuronalny jest złożony z neuronów, które są komórkami biologicznymi wyspecjalizowanymi do przetwarzania i przekazywania sygnałów. Punktem wyjścia wszelkich modeli neuronu są wyniki badań neurofizjologów, mające już swoją historię. Twórcą teorii neuronalnej jest hiszpański neurohistolog i neuroanatom Santiago Ramón y Cajal (1852–1934), który „w uznaniu za pracę nad strukturą układu nerwowego” otrzymał w 1906 roku, wraz

z Camillem Golgim (1843–1926), Nagrodę Nobla w dziedzinie fizjologii lub medycyny. Przed odkryciami obu uczonych mózg uważano za jednolitą masę. Metoda barwienia preparatów histologicznych, opracowana przez Golgiego w 1873 roku, pozwoliła ujrzeć strukturę mózgu. Golgi jednak uważał, że mózg jest strukturą będącą siecią kanalików. Ramón y Cajal na podstawie własnych obserwacji, wykorzystując metodę Golgiego, doszedł do wniosku, że mózg składa się z odrębnych komórek, które się nie stykają, w szczególności protoplazma poszczególnych neuronów nie miesza się ze sobą. Z kolei angielski lekarz i naukowiec Richard Caton (1842–1926) w roku 1875 poinformował o swoich odkryciach dotyczących elektrycznej aktywności mózgu.

Istotny wkład do badań nad układem nerwowym wniósł angielski lekarz i fizjolog Charles Sherrington (1857–1952), laureat Nagrody Nobla, wraz z Edgarem Adrianem (1889–1977), w dziedzinie fizjologii lub medycyny w 1932 roku „za odkrycia dotyczące funkcji neuronów”. W roku 1887 Sherrington wprowadził pojęcie synapsy jako połączenia między dwoma neuronami lub między neuronem a komórką pobudliwą innego typu, na przykład mięśniową. Synapsa składa się z trzech części – kolbki presynaptycznej, czyli zakończenia neuronu przekazującego sygnał, tzw. neuronu presynaptycznego, szczeliny synaptycznej między neuronem presynaptycznym a komórką odbierającą sygnał, tzw. komórką postsynaptyczną oraz z receptora komórki postsynaptycznej. Rozróżniamy dwa typy synaps – chemiczną i elektryczną. W synapsie chemicznej sygnał między komórkami jest przekazywany za pośrednictwem elektrycznie obojętnej substancji chemicznych – neurotransmiterów. Szerokość szczeliny synaptycznej wynosi ok. 20 nm. W synapsie elektrycznej szerokość szczeliny synaptycznej wynosi ok. 4 nm, a sygnał jest przekazywany poprzez transport jonów. W synapsach elektrycznych szybkość przekazywania sygnału jest prawie tysiąc razy większa niż w synapsach chemicznych, ale sygnał jest słabszy, a możliwość sterowania sygnałem – znikoma. Ponadto sygnalizacja jest dwukierunkowa. Działanie chemicznej synapsy wyjaśnili dwaj uczeni – brytyjski fizjolog i farmakolog, Henry Dale (1875–1968) oraz austriacki farmakolog Otto Loewi (1873–1961). Obaj zostali uhonorowani Nagrodą Nobla w 1936 roku za „odkrycie dotyczące chemicznego przekazywania impulsów nerwowych”. Angielscy naukowcy, biochemik Alan Hodgkin (1914–1998) oraz fizjolog i biofizyk Andrew Huxley (1917–2012), opisali mechanizmy

transportu jonów w neuronach, który jest podstawą generacji potencjałów czynnościowych. Opracowali także zarówno matematyczny model przewodnictwa jonowego w neuronie, bazujący na układzie równań różniczkowych zwyczajnych, oraz funkcjonalnie równoważny model elektroniczny (Hodgkin & Huxley, 1952). Utorowało to drogę do matematycznych i elektronicznych modeli zjawisk neurodynamicznych. Modele tego typu były później wykorzystywane jako baza wielu typów sztucznych sieci neuronowych. Za swoje badania otrzymali Nagrodę Nobla z dziedziny fizjologii lub medycyny w 1963 roku.

Drugim, obok neuronów, składnikiem tkanki nerwowej są komórki glejowe, odkryte w 1851 roku przez niemieckiego anatoma Heinricha Müllera (1820–1864). Pod koniec lat 50. XIX wieku niemiecki patolog Rudolf Virchow (1821–1902), korzystając z techniki opracowanej przez Golgiego, odkrył astrocyty – jeden z rodzajów komórek glejowych. Do końca XX wieku uważano, że komórki glejowe tworzą mielinowe osłonki aksonów, czyli pojedynczej wypustki neuronalnej, która przewodzi impuls do synapsy, oraz przestrzenne rusztowanie, na którym rozpięte są neurony. W XXI wieku jednak odkryto, że przynajmniej niektóre rodzaje komórek glejowych biorą udział w przewodzeniu sygnałów, wymieniając sygnały z neuronami (Fields, 2011). Jak dotąd jednak komórki glejowe nie występują w cybernetycznych modelach układu nerwowego ani, co za tym idzie, nie są uwzględniane w sztucznych sieciach neuronowych, które – jak dotychczas – są oparte wyłącznie na modelach synapsy chemicznej.

Jak już wspomniano, pierwszymi dwoma krokami w rozwoju sztucznych sieci neuronowych było opracowanie progowego modelu neuronu McCullocha i Pittsa w 1943 roku oraz implementacja perceptronu w roku 1957, zbudowanego ze sztucznych neuronów opartych na tym modelu. Był on przeznaczony do rozpoznawania znaków alfanumerycznych, jednak okazał się wrażliwy na skalowanie rozpoznawanego wzorca. Dla jednowarstwowego perceptronu Rosenblatt opracował algorytm iteracyjnego ustawiania wag (algorytm uczenia sieci) i dowiódł jego zbieżności. Podobną konstrukcją była sieć ADALINE, zbudowana w 1959 roku przez Bernarda Widrowa (ur. 1929) i Tedda Hoffa (ur. 1937). Sieć ta była strukturą jednowarstwową, zaimplementowaną jako układ elektroniczny (Widrow & Hoff, 1960). Przeznaczona była do adaptacyjnego przetwarzania sygnałów.

Marvin Minsky (1927–2016) w drugiej połowie lat 60. XX wieku przeanalizował możliwości znanych wtedy sztucznych sieci neuronowych i poddał je krytyce, wskazując na ich istotne ograniczenia. Podał przykłady problemów, których nie da się rozwiązać przy pomocy sieci jednowarstwowych. Przedyskutował również naukę sieci składających się z więcej niż jednej warstwy (Minsky & Papert, 1969), wskazując na fundamentalne problemy związane z nauką tzw. warstw ukrytych, czyli tych, które nie przesyłają sygnałów bezpośrednio na wyjście całego układu, tylko do innych warstw sieci. Ze swoich analiz wysnuł wniosek, że sztuczne sieci neuronowe nie są obiecującym kierunkiem w rozwoju systemów sztucznej inteligencji.

Po początkowym dużym zainteresowaniu sztucznymi sieciami neuronowymi w drugiej połowie lat 50. i w latach 60. pod wpływem książki Minsky'ego i Paperta w latach 70. zainteresowanie problematyką badawczą związaną z sieciami neuronowymi znacząco osłabło. Mimo tego w latach 70. odnotowano kilka interesujących wyników. Jednym z tych, które istotnie przyczyniły się do przełamania impasu spowodowanego krytyką Minsky'ego, było zaproponowanie nauki sieci wielowarstwowej przy pomocy algorytmu wstecznej propagacji błędu, opracowanego w 1974 roku (Werbos, 1974). Podstawą sukcesu było staranne opracowanie formalnego modelu pewnego sposobu nauki, opartego na tym, że uczący się układ cybernetyczny dysponuje wzorcami wyników prawidłowych w danym kontekście. W przypadku sztucznych sieci neuronowych kontekstem tym jest wektor sygnału podawanego na wejście sieci, a wynikiem – wektor sygnału wyjściowego sieci. Zbiór par składających się z wektora sygnału wejściowego i odpowiadającego mu wektora sygnału wyjściowego, który jest poprawną reakcją sieci na dany sygnał wejściowy, tworzy tzw. ciąg uczący, czyli zbiór sygnałów wejściowych, dla których znamy poprawną odpowiedź układu i zbiór odpowiedzi, który będzie punktem odniesienia dla algorytmu zmiany wag. Taka idea nauki sieci należy do grupy algorytmów „z nauczycielem” (patrz niżej). Mając miarę poprawności sieci, możemy opracować taki algorytm zmiany wag neuronów, żeby sieć działała coraz lepiej. Jedną z najprostszyc miar tego typu jest suma kwadratów różnic wektorów poprawnych odpowiedzi danych w ciągu uczącym i wektorów odpowiedzi sieci na dany bodziec z ciągu uczącego, przy czym przez kwadrat wektora należy rozumieć iloczyn skalarny wektora przez siebie. Ponieważ jest on

liczbą nieujemną, równą zero tylko wtedy, gdy wszystkie wektory aktualnych odpowiedzi sieci są równe wektorom odpowiedzi poprawnych, miara ta będzie równa zero wyłącznie w tym wypadku, gdy sieć będzie działała absolutnie poprawnie. Należy więc zastosować taki algorytm modyfikacji wag, aby wspomniana miara, będąca funkcją wszystkich wag w sieci, z każdym krokiem nauki malała. Ponieważ zastosowana miara jest funkcją o wartościach rzeczywistych i argumentach będących wagami sieci, można zastosować algorytmy gradientowe do znalezienia minimum tej funkcji. Najprostszym tego typu algorytmem jest algorytm największego spadku gradientu, znany od połowy XIX wieku. Jego zastosowanie w kontekście nauki sieci warstwowej z opisaną wyżej kwadratową funkcją błędu daje algorytm wstecznej propagacji błędu.

Chcąc wykorzystać metody gradientowe do nauki sieci warstwowych, należało nieco zmodyfikować model McCullocha-Pittsa, który jako model progowy z funkcją skoku jednostkowego nie był różniczkowalny. W związku z tym wprowadzono różniczkowalną funkcję aktywności neuronu. Zgodnie z tym zmodyfikowanym modelem całkowite wzbudzenie neuronu jest przekształcane przy pomocy funkcji różniczkowalnej. Wynik tego przekształcenia to sygnał wyjściowy neuronu. Metoda największego spadku gradientu jest, w grupie metod gradientowych, stosunkowo powolna. Dlatego dziś stosuje się raczej szybsze metody gradientowe. Tak więc zaproponowanie adekwatnego modelu procesu nauki pozwoliło zastosować do trenowania sieci neuronowych opracowaną dużo wcześniej grupę metod numerycznych. Metoda wstecznej propagacji błędu nie od razu stała się znaną i stosowaną metodą. Dopiero kilka kolejnych publikacji, których autorzy opracowali ją niezależnie od wyniku opisanego w latach 70., spowodowało, że w latach 80. zaczęto ją stosować powszechnie. Spowodowało to ponowny wzrost zainteresowania sztucznymi sieciami neuronowymi i w konsekwencji zaowocowało stworzeniem nowych modeli neuronu, nowych typów sieci neuronowych i zaproponowaniem nowych algorytmów nauki.

Fiński informatyk Teuvo Kohonen (1934–2021) zaproponował nowy rodzaj sieci, nazywanych dziś sieciami Kohonena (Kohonen, 1982). Są one modelem pamięci skojarzeniowej i uczy się ich przy wykorzystaniu idei map samoorganizujących się (jest to również alternatywna nazwa sieci Kohonena), trenowanych przy użyciu uczenia nienadzorowanego. Ten algorytm uczenia polega na grupowaniu sygnałów wejściowych,

a jedynym kryterium nauki jest ich podobieństwo. W wyniku nauki wagi neuronów ustawione są w ten sposób, że podobne sygnały pobudzają te same neurony.

Kolejnym typem sieci neuronowej jest sieć Hopfielda, zaproponowana przez Johna Hopfielda (ur. 1933), amerykańskiego fizyka, który zajmował się także sieciami neuronowymi jako układami obliczeniowymi (Hopfield, 1982). Sieć Hopfielda składa się z neuronów typu McCullocha-Pittsa połączonych „każdy z każdym”. W związku z tym po wzbudzeniu niektórych lub wszystkich neuronów w sieci sygnały w sieci krążą, ponieważ połączenia neuronów mają strukturę pętli. Odpowiedzią sieci jest stan końcowy iteracyjnego procesu pobudzeń neuronu. Jako system sztucznej inteligencji sieć Hopfielda jest używana jako model pamięci skojarzeniowej oraz jako system optymalizacyjny (Mańdziuk, 2000). Sieci typu Hopfielda są również szeroko stosowane jako systemy modelujące dynamikę przebiegu pobudzeń w biologicznych systemach neuronalnych. Sieci komórkowe, wprowadzone w 1988 roku, są to sieci, które tworzą strukturę połączeń lokalnych. Realizowana jest ona w ten sposób, że neurony umieszczamy w węzłach pewnej siatki (najczęściej jest to siatka kwadratowa na płaszczyźnie, ale możliwe są też inne struktury) i każdy neuron jest połączony jedynie z neuronami leżącymi w jego sąsiedztwie, najczęściej bezpośrednim, rzadziej w sąsiedztwie o zadanym promieniu. Sąsiedztwo należy tu rozumieć w sensie zgodnym ze strukturą sieci. Neurony leżące na brzegu sieci naruszają symetrię połączeń międzyneuronalnych. W związku z tym często łączy się ze sobą przeciwległe krawędzie sieci. Od strony geometrycznej oznacza to, że sieć, zamiast na płaszczyźnie, rozpinamy na powierzchni torusa. Dynamikę zmian stanów sieci komórkowej opisują równania różniczkowe dla sieci zbudowanych z neuronów analogowych lub równania różnicowe w przeciwnym wypadku. Dynamikę stanu pojedynczego neuronu można opisać przy pomocy równania różniczkowego zwyczajnego, natomiast dynamikę całej sieci opisuje równanie różniczkowe cząstkowe. Dlatego też sieci komórkowe stosowane są również do numerycznego rozwiązywania równań różniczkowych cząstkowych oraz do ich symulacji z czasem ciągłym w przypadku zastosowania sieci komórkowych zaimplementowanych przy pomocy elektroniki analogowej.

Od lat 80. XX wieku liczba nowych rodzajów sieci gwałtownie wzrasta. Proponowane są zarówno nowe typy sieci, jak też nowe modele

neuronów. Ponadto implementuje się wielomodułowe systemy hybrydowe, w których sieć neuronowa jest jednym z modułów. Obszary zastosowań sieci neuronowych są bardzo liczne.

Problemowe ujęcie pojęcia i obszaru badań

Jak już wspomniano, jednym z zasadniczych zagadnień związanych z problematyką sztucznych sieci neuronowych jest proces ich nauki, czyli algorytm ustawienia parametrów poszczególnych neuronów. Opracowano wiele takich algorytmów i choćby wymienienie wszystkich wykraczałoby poza ramy niniejszego opracowania. Przeprowadźmy więc taksonomię metodyki nauki sieci neuronowych na najwyższym poziomie ogólności.

Rzadko się zdarza, że potrafimy obliczyć parametry sztucznej sieci neuronowej, analizując od strony teoretycznej specyfikę sieci oraz problem, który sieć ma rozwiązywać. Taka sytuacja ma miejsce w przypadku małych sieci i prostych problemów. Czasem rezultat może być osiągnięty przy pomocy zaawansowanego aparatu matematycznego, adaptacji modelu fizycznego lub analizy jakiegoś algorytmu nauki sieci, który może dla danego typu sieci dać wzór algebraiczny na ustawienie wag w jednym akcie doboru. Taka sytuacja zachodzi na przykład dla niektórych problemów, do których stosujemy sieci Hopfielda. Na ogół jednak uczenie sieci neuronowych ma charakter procesu iteracyjnego, w którym parametry neuronów zmieniane są krok po kroku, za każdym razem o pewną niewielką wartość.

Istnieją dwie zasadnicze grupy iteracyjnych metod nauki sieci. Pierwszą z nich są algorytmy uczenia z nadzorem, zwane też uczeniem z nauczycielem (ang. *supervised learning*, *supervised training*). Mogą być one stosowane wtedy, gdy mamy pewne informacje dotyczące poprawnego działania nauczonej sieci. Informacje te wykorzystujemy podczas uczenia sieci neuronowej. W zależności od tego, jaki mają one charakter i w jaki sposób je wykorzystujemy, możemy w metodach z nadzorem wyróżnić różne strategie. W metodach tych dysponujemy zbiorem sygnałów wejściowych, dla których znamy prawidłowe odpowiedzi, jakie sieć powinna generować na wyjściu. Ciąg złożony z par wektorów, w których pierwszy jest wektorem sygnału wejściowego, a drugi wektorem kodującym poprawną odpowiedź sieci, konstytuuje tzw.

ciąg uczący. Algorytm uczenia polega na porównaniu znanej poprawnej odpowiedzi z faktyczną odpowiedzią sieci oraz zmianie parametrów sieci w taki sposób, aby podczas następnej prezentacji tego samego sygnału wejściowego odpowiedź sieci była bliższa poprawnej. Bardzo często stosowanymi metodami tego typu są wspomniane wyżej metody gradientowe, w których znany i ustalony ciąg uczący jest podstawą zdefiniowania różniczkowalnej funkcji błędu posiadającej minimum globalne dla tych wag sieci, dla których odpowiedzi sieci pokrywają się z odpowiedziami prawidłowymi. Metody gradientowe są metodami numerycznymi rozwiązywania równań różniczkowych zwyczajnych lub pewnych problemów optymalizacyjnych zaadaptowanymi do nauki sieci neuronowych. Metody te są deterministyczne, co oznacza, że w każdym kroku nauki aktualne parametry sieci jednoznacznie określają ich zmianę. Metody stochastyczne natomiast polegają na losowym ustawianiu parametrów sieci w każdym kroku nauki zgodnie z rozkładami prawdopodobieństwa, które definiujemy na podstawie informacji posiadanych o rozwiązaniu. Przykładem tego typu strategii są algorytmy genetyczne i strategie ewolucyjne, które też mogą być wykorzystywane do nauki sieci neuronowych.

Drugą zasadniczą grupą nauki sieci neuronowych są metody uczenia bez nadzoru. Charakteryzują się one tym, że nie mamy żadnej informacji o ewentualnym rozwiązaniu problemu. Jedyne więc, co sieć może zrobić, to zaklasyfikować sygnały podobne do siebie do tej samej klasy sygnałów wejściowych. Z matematycznego punktu widzenia tak postawiony problem wyznacza pewną relację równoważności – dwa sygnały wejściowe są ze sobą w relacji, jeśli są do siebie podobne. Oczywiście dla danego problemu wzmiankowane podobieństwo musi być dobrze zdefiniowane. Wynikiem uczenia sieci będzie podział przestrzeni sygnałów wejściowych na klasy równoważności. Te metody dzielą się na dwie zasadnicze grupy. Pierwszą z nich są metody typu konkurencyjnego. Polegają one na tym, że po zaprezentowaniu sygnału wejściowego znajdowany jest ten neuron, którego wektor wag jest najbardziej podobny do zaprezentowanego sygnału, i modyfikowane są wagi tylko tego neuronu w ten sposób, aby oba wektory były jeszcze bardziej do siebie podobne (algorytmy typu *winner takes all*), lub modyfikujemy wagi kilku najbardziej podobnych do zaprezentowanego sygnału neuronów (algorytmy typu *winner takes most*). Przy pomocy tej metody trenowane są

sieci Kohonena (Tadeusiewicz, 1993). Drugą grupą metod nauki bez nadzoru są metody oparte na regule Donalda Hebba (1904–1985), kanadyjskiego neuropsychologa, który badał związki między funkcjami psychicznymi a strukturą i aktywnością mózgu. Jego obserwacje i eksperymenty doprowadziły go do wniosku, że wielokrotnie pobudzone połączenie synaptyczne między dwoma neuronami zwiększa swoją przewodność (Hebb, 1949).

Kolejną ważną tematyką związaną ze sztucznymi sieciami neuronowymi są ich aspekty matematyczne (Bielecki, 2019). Zasadniczy problem badawczy stanowią matematyczne własności procesu nauki, dynamiczne własności nauczonych sieci rekurencyjnych oraz możliwości obliczeniowe sieci neuronowych. Przy analizie tych problemów często adaptuje się modele matematyczne opracowane do opisu zjawisk fizycznych (Hertz et al., 1991). Jak już wspomniano, algorytmy nauki sieci neuronowych w większości przypadków mają charakter iteracyjny i są implementacją procesu numerycznego. W tym kontekście pojawiają się pytania o zbieżność i stabilność procesu nauki. W przypadku gdy teoretycznym modelem procesu nauki jest równanie różniczkowe, pojawia się dodatkowy problem, na ile dynamika generowana przez dyskretyzację danego równania różniczkowego poprzez ustalenie kroku czasowego przenosi się na dyskretny układ dynamiczny generowany przez metodę numeryczną zastosowaną do danego równania różniczkowego. Problemy te są związane z topologicznym sprzężeniem dyskretnych układów dynamicznych oraz z tzw. własnością cieniowania (ang. *shadowing property*, zwana też własnością *pseudo-orbit tracing property*). Kolejnym matematycznym problemem jest wykorzystanie sieci jako reprezentacji odwzorowania wielowymiarowego, dla którego znamy zależności w pewnej liczbie punktów danych w ciągu uczącym. W drugiej połowie lat 80. i w latach 90. prowadzono intensywne badania teoretyczne w tym zakresie. Wychodząc od znanych wyników matematycznych, przede wszystkim od twierdzenia Kołmogorowa o aproksymacji funkcji wielu zmiennych przy pomocy rodziny funkcji jednej zmiennej, wykazano, że warstwowa sieć neuronowa z jedną warstwą ukrytą i neuronami z logistyczną funkcją aktywacji w sieci ukrytej i liniową funkcją aktywacji w warstwie wyjściowej może na zbiorze zwartym przybliżać z dowolnie małym, ale niezerowym błędem każdą funkcję ciągłą. Oznacza to, że taka sieć jest uniwersalnym aproksymatorem.

Istotną tematyką badawczą, będącą kontynuacją badań związanych z początkiem sztucznych sieci neuronowych, jest ich wykorzystanie jako narzędzi do modelowania struktur i procesów na poziomie neurofizjologicznym i, co za tym idzie, procesów psychicznych, zarówno w normie, jak i w patologii (Duch, 2000). Jak już wspomniano, przy pomocy sieci neuronowych można modelować pamięć asocjacyjną oraz badać jej odporność na uszkodzenia poprzez sprawdzenie, jak usunięcie pewnej liczby neuronów wpływa na jej zawartość. Z kolei uszkodzenia dróg czuciowych powodują zmianę organizacji map kory mózgu, co może być modelowane przy pomocy map samoorganizujących się typu Kohonena. Modele neuronalne pomogły również wyjaśnić, jak zaburzenia hamowania pobudzenia neuronów mogą doprowadzić do głębokich zaburzeń pamięci. Obecność pasożytniczych atraktorów w sieciach rekurencyjnych pozwoliła z kolei rzucić światło na takie zaburzenia psychiczne jak halucynacje i urojenia. Badania tego typu są intensywnie rozwijane.

Sieci neuronowe znajdują wielorakie zastosowania jako systemy sztucznej inteligencji. Przykładowo w technice i przemyśle sieci neuronowe stosuje się do sterowania oraz do identyfikacji obiektów dynamicznych. W tym ostatnim przypadku częstym zadaniem jest znalezienie nieliniowego odwzorowania aproksymującego dany układ techniczny. Przy pomocy warstwowych sieci neuronowych, będących uniwersalnymi aproksymatorami, można problem nieliniowej identyfikacji skutecznie rozwiązać. Sieci neuronowe stosuje się również do monitorowania i diagnozowania urządzeń technicznych, zwłaszcza o dużym stopniu złożoności. Typowa metoda polega na budowie neuronalnego klasyfikatora stanów urządzenia na podstawie danych pomiarowych. Klasyfikatory neuronalne są układami mogącymi działać w czasie rzeczywistym. W medycynie sieci neuronowe stosuje się do rozpoznawania obrazów medycznych oraz do wspomagania procesu diagnostyki. W ekonomii sieci neuronowe używane są do predykcji sytuacji na rynku, na przykład do predykcji poziomu popytu i podaży na dane towary i usługi. Duże znaczenie ma predykcja zapotrzebowania na energię elektryczną w związku z faktem, że wciąż są istotne trudności z magazynowaniem energii elektrycznej na skalę przemysłową. Dodatkową trudnością jest inwestycja w odnawialne źródła energii, które, zwłaszcza w przypadku elektrowni wiatrowych, wykazują silną zależność możliwości produkcyjnych od warunków atmosferycznych. Sieci neuronowe stosowane są

również do wspomagania decyzji oraz w detekcji rozpoznawania i klasyfikacji zachowań rynków.

Polscy naukowcy wnoszą cenny wkład do badań nad sztucznymi sieciami neuronowymi zarówno w aspekcie teoretycznym, jak i aplikacyjnym. Działający od 1979 w Stanach Zjednoczonych Jacek Żurada (ur. 1944) prowadzi badania nad algorytmami nauki sieci warstwowych i warunkami stabilności sieci. Ryszard Tadeusiewicz (ur. 1947), przez całą swoją karierę akademicką związany z Akademią Górniczo-Hutniczą, zajmuje się modelowaniem struktur układu nerwowego, zastosowaniem sieci neuronowych w klasyfikacji i analizowaniu obrazów, w szczególności wyników obrazowania medycznego oraz zastosowaniem sieci neuronowych w przemyśle. Józef Korbicz (ur. 1951) zajmuje się głównie zastosowaniami sieci neuronowych w technice i przemyśle. Leszek Rutkowski (ur. 1952) zajmuje się teoretycznymi podstawami sieci neuronowych, w szczególności neuro-rozmytymi systemami hybrydowymi oraz zastosowaniami do rozpoznawania obrazów i do sterowania. Włodzisław Duch (ur. 1954) bada zastosowania sieci neuronowych do modelowania procesów neurologicznych i psychicznych zarówno w normie, jak i w patologii (Duch, 2000).

Refleksja systematyczna i wnioski

Sztuczne sieci neuronowe są specyficznymi systemami cybernetycznymi. Istotną ich cechą jest, że posiadają dobrze zdefiniowany pierwowzór biologiczny. Pierwowzór ten jednakże to układ o tak dużym stopniu złożoności, że na obecnym etapie rozwoju informatyki i elektroniki jesteśmy w stanie modelować jedynie niewielką część aspektów systemu nerwowego, tym bardziej że mimo postępu neurofizjologii nasza wiedza zarówno o układzie nerwowym jako całości i o jego poszczególnych modułach – także o pojedynczym neuronie – jest dalece niekompletna. Dodatkową trudnością jest nie do końca wyjaśniona funkcja komórek glejowych. Pomimo powyższych zastrzeżeń modelowanie struktury i funkcji układu nerwowego, jak również funkcjonalności psychicznych wniosło istotny przyczynek do neurofizjologii i psychiatrii. Należy oczekiwać dalszej eksploatacji tego kierunku badań naukowych. Trzeba jednak wyraźnie zaznaczyć, że wielu naukowców krytycznie

wypowiada się o możliwości implementacji sztucznego układu inteligentnego, w szczególności opartego na sztucznych sieciach neuronowych, którego całościowe możliwości byłyby porównywalne z możliwościami ludzkiej inteligencji, przede wszystkim w jej szeroko rozumianym aspekcie twórczym. Z drugiej strony, niezależnie od biologicznych korzeni, sztuczne sieci neuronowe stały się jednym z typów systemów sztucznej inteligencji, które są używane do rozwiązywania szerokiej klasy zadań. Ich istotną wadą jest kłopotliwy proces nauki oraz fakt, że należą do grupy systemów obliczeniowych, w których wiedza nie jest reprezentowana w sposób jawny. Nie pozwala to śledzić procesu rozumowania, który łączy fakty podane na wejście systemu z odpowiedzią systemu. Ich niewątpliwą zaletą jest to, że wiedza, jaką musimy dysponować, żeby przystąpić do rozwiązywania problemu, jest niewielka. Wystarczy dobrać wystarczającą liczbę reprezentatywnych przykładów do ciągu uczącego.

Systemy sztucznej inteligencji, w tym sieci neuronowe, są systemami pozwalającymi na szybką analizę dużej ilości danych różnego typu oraz wprowadzania systemów automatycznych o coraz większym stopniu autonomii. Z jednej strony daje to wielkie możliwości rozwoju techniki i rozwiązywania kolejnych problemów cywilizacyjnych. Z drugiej strony staje się groźnym narzędziem umożliwiającym inwigilację jednostki i kontrolowanie społeczeństwa.

BIBLIOGRAFIA

- Bielecki, A. (2019). Models of Neurons and Perseptrons: Selected Problems and Challenges. *Studies in Computational Intelligence*, 770. Cham: Springer.
- Duch, W. (2000). Sieci neuronowe w modelowaniu zaburzeń neuropsychologicznych i chorób psychicznych. W: W. Duch, J. Korbicz, L. Rutkowski, R. Tadeusiewicz (Red.), *Sieci neuronowe* (s. 589–616). Warszawa: Akademicka Oficyna Wydawnicza EXIT.
- Hebb, D.O. (1949). *The Organization of Behaviour, a Neuropsychological Theory*. New York: Wiley.
- Fields R.D. (2011). *Drugi mózg*. Warszawa: Prószyński i S-ka.
- Hertz, J., Krogh, A., Palmer, R.G. (1991). *Wstęp do teorii obliczeń neuronowych*. Warszawa: Wydawnictwa Naukowo-Techniczne.

- Hodgkin, A.L., & Huxley, A.F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117(4), 500–544.
- Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8), 2554–2558.
- Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59–69.
- Mańdziuk, J. (2000). *Sieci neuronowe typu Hopfielda*. Warszawa: Akademicka Oficyna Wydawnicza EXIT.
- McCulloch, W.S., & Pitts, W. (1943). A logical calculus for the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–133.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge: MIT Press.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.
- Tadeusiewicz, R. (1993). *Sieci neuronowe*. Warszawa: Akademicka Oficyna Wydawnicza.
- Werbos, P. (1974). *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Systems*. Ph.D. Thesis. Cambridge: Harvard University.
- Widrow, B., & Hoff, M.E. (1960). Adaptive switching circuits. W: 1960 *IRE WESCON Convention Record*, 4 (s. 96–104).

Mariusz Flasiński

Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0002-0910-4928>

<https://doi.org/10.35765/slowniki.491>

Sztuczna inteligencja

Streszczenie

DEFINICJA POJĘCIA: W rozdziale przedstawione zostały różne definicje pojęcia „sztuczna inteligencja” oraz zamieszczona została typologia systemów sztucznej inteligencji.

ANALIZA HISTORYCZNA POJĘCIA: Rozdział zawiera chronologiczny przegląd podstawowych paradygmatów sztucznej inteligencji (AI) począwszy od lat 50. XX wieku. W ramach podobszaru symbolicznej sztucznej inteligencji omówione zostały takie podejścia do konstrukcji systemów AI jak: symulacja kognitywna, podejście oparte na logice oraz podejście oparte na wiedzy. W dalszej części rozdziału przedstawione zostały następujące podejścia w podobszarze inteligencji obliczeniowej: modele inspirowane matematyką, modele konekjonistyczne, oraz obliczenia ewolucyjne.

UJĘCIE PROBLEMOWE POJĘCIA: W rozdziale zostały scharakteryzowane najważniejsze metody/modely sztucznej inteligencji: metody przeszukiwania przestrzeni stanów, modele logiki matematycznej, modele oparte na wiedzy regułowej, strukturalne modele reprezentacji wiedzy, syntaktyczna analiza wzorców, rozpoznawanie obrazów, algorytmy genetyczne, programowanie genetyczne i sztuczne sieci neuronowe.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Ostatni rozdział zawiera dyskusję poświęconą fundamentalnym barierom metodologicznym w konstrukcji systemów „inteligentnych”. W tym celu wprowadzone zostały dychotomie: silna AI vs słaba AI oraz ogólna AI (AGI) vs specyficzna AI, a także pojęcie sztucznej superinteligencji.

Słowa kluczowe: sztuczna inteligencja (AI), symboliczna sztuczna inteligencja, inteligencja obliczeniowa, silna sztuczna inteligencja, słaba sztuczna inteligencja

Definicja pojęcia

W literaturze przedmiotu można znaleźć wiele różniących się definicji pojęcia „sztuczna inteligencja” (ang. *Artificial Intelligence*, AI). Wynika to tak z historycznych zaszłości, jak i z braku konsensusu odnośnie do zakresu i przedmiotu tego pojęcia. Poniżej przedstawimy dwie definicje zamieszczone w podstawowych podręcznikach wprowadzających do problematyki sztucznej inteligencji. Pierwsza z nich jest próbą określenia sensu (istoty) pojęcia „sztuczna inteligencja”, podczas gdy druga definicja próbuje odnieść się do różnych definicji, które pojawiały się w literaturze przedmiotu na przestrzeni wielu lat w związku z rozwojem tego obszaru badań, a następnie próbuje dokonać ich uogólnienia.

Według *Introduction to Artificial Intelligence* (Fłasiński, 2016) termin „sztuczna inteligencja” ma dwa podstawowe znaczenia. Zgodnie z pierwszym znaczeniem oznacza wspólny obszar informatyki i robotyki, w którym prowadzone są badania nad konstrukcją systemów (informatycznych i/lub robotyki) wykonujących czynności wymagające w przypadku człowieka inteligencji (czasami systemy takie nazywane są inteligentnymi). W drugim znaczeniu „sztuczna inteligencja” oznacza cechę sztucznych systemów, która umożliwia im działania w przypadku człowieka wymagające inteligencji. Jak więc widać, w definicji tej odwołano się do pojęcia inteligencji człowieka, zostawiając jego doprecyzowanie dyscyplinom naukowym, dla których jest ono przedmiotem badań (filozofia, psychologia, kognitywistyka).

Russell i Norvig (2020) poprzedzają wprowadzenie definicji sztucznej inteligencji spostrzeżeniem, że można wprowadzić taksonomię definicji systemów sztucznej inteligencji, które pojawiały się na przestrzeni wielu lat w literaturze przedmiotu, ze względu na dwa kryteria. Pierwsze z nich dotyczy tego, czy skupiamy się na (inteligentnym) zachowaniu się tych systemów czy też na procesach umysłowo-poznawczych symulowanych w tych systemach. Drugie kryterium bierze pod uwagę fakt, czy jako punkt odniesienia zachowania się systemu/procesów umysłowo-poznawczych przyjmujemy człowieka, czy też racjonalność rozumianą tutaj jako „idealną” inteligencję (w działaniu). Na podstawie tak określonych kryteriów otrzymujemy następującą typologię systemów sztucznej inteligencji:

- systemy, które myślą tak, jak ludzie – np. definicja w: Haugeland, 1985,

- systemy, które zachowują się tak, jak ludzie – np. definicja w: Rich & Knight, 1991,
- systemy, które myślą racjonalnie – np. definicja w: Winston, 1992,
- systemy, które zachowują się racjonalnie – np. definicja w: Nilsson, 1998.

W celu poradzenia sobie z taką różnorodnością definicji Russell i Norvig wprowadzają pojęcie „sztucznej inteligencji”, opierając się na informatycznej koncepcji agenta. Agent jest zdefiniowany jako system, który postrzega środowisko (w którym się znajduje) za pomocą sensorów i oddziałuje (autonomicznie) na to środowisko poprzez efekторы w celu osiągnięcia zamierzonych celów. Jeśli agent potrafi udoskonalać swoje zachowanie poprzez uczenie się i/lub pozyskiwanie wiedzy, to mówimy wtedy o inteligentnym agencie. Zauważmy, że pod taką definicję agenta podpadają także ludzie, roboty, chatboty itp. W rezultacie sztuczna inteligencja jest zdefiniowana u Russella i Norviga (2020, s. viii) jako badania nad konstrukcją inteligentnych agentów.

Analiza historyczna pojęcia

Analiza historyczna przedstawiona w niniejszym rozdziale opiera się na faktach podanych w książce Flasińskiego (2016). Początki sztucznej inteligencji jako obszaru badań nad konstrukcją systemów inteligentnych sięgają połowy lat 50. ubiegłego wieku. W 1955 roku Allen Newell i Herbert Simon w ramach badań prowadzonych w Carnegie-Mellon University skonstruowali system Logic Theorist, który potrafił wygenerować dowody matematyczne blisko czterdziestu początkowych twierdzeń z fundamentalnego dzieła Alfreda N. Whiteheada i Bertranda Russella *Principia Mathematica*. Logic Theorist można uznać za pierwszy system sztucznej inteligencji, a rok 1955 – za początek badań w tym obszarze. Równocześnie rok ten można uznać za moment pojawienia się w obszarze sztucznej inteligencji jej pierwszego podejścia paradygmatycznego nazywanego symulacją kognitywną. Podstawowe założenie tego podejścia mówi, że systemy inteligentne należy konstruować jako symulatory ludzkich procesów umysłowo-poznawczych.

Rok później, w 1956 roku, na konferencji w Dartmouth College poświęconej „mechanizacji procesów myślowych” John McCarthy z Massachusetts

Institute of Technology zaproponował inne ważne podejście paradygmatyczne, a mianowicie podejście oparte na logice matematycznej. Założeniem tego podejścia jest konstrukcja systemów inteligentnych opartych na różnych modelach logiki matematycznej. Sam McCarthy zasugerował wykorzystanie do tego celu rachunku lambda zdefiniowanego w latach 30. XX wieku przez Alonza Churcha i Stephena C. Kleene'ego, a także skonstruował na bazie tego rachunku współcześnie standardowo używany w sztucznej inteligencji język programowania LISP. W 1973 roku Alain Colmerauer i Philippe Roussel z Université d'Aix-Marseille uzupełnili to podejście, wprowadzając – jako model symulacji rozumowania w systemach sztucznej inteligencji – logikę pierwszego rzędu z wnioskowaniem na podstawie metody rezolucji, a także konstruując drugi standardowo używany język programowania w sztucznej inteligencji, mianowicie PROLOG.

Wspólną cechą obu podejść omówionych powyżej jest to, że próbuje się za ich pomocą symulować inteligencję jako taką, jakby to można było określić – inteligencję ogólną. Natomiast badania prowadzone na Stanford University w połowie lat 60. ubiegłego wieku przez Edwarda Feigenbauma i Joshuę Lederberga miały mniej ambitny cel – ukierunkowane były bowiem na konstrukcję systemu rozwiązującego skomplikowane problemy w dobrze określonym obszarze zastosowań (w przypadku tych badań była to identyfikacja struktury molekularnej nieznanych związków organicznych na podstawie analizy widm spektroskopowych, wykonywana przez skonstruowany przez nich system DENDRAL). Systemy takie konstruowane są na podstawie bazy wiedzy specjalistycznej pozyskanej od ekspertów w danej dziedzinie problemowej i dlatego nazywa się je systemami eksperckimi, a samo podejście określa się jako podejście oparte na wiedzy (dziedzinowej). Istnieją dwa podejścia w sztucznej inteligencji, ważne ze względu na ogromny zakres ich praktycznych zastosowań, które można potraktować jako rozwinięcia tego paradygmatu. Są to: syntaktyczne (strukturalne) rozpoznawanie wzorców (Flasiński, 2019), oparte na koncepcji gramatyk generatywnych Noama Chomsky'ego, mające zastosowanie m.in. w analizie obrazów i wizji komputerowej, bioinformatyce, automatycznym rozpoznawaniu/identyfikacji sygnałów (medycznych (np. EKG, EEG), przemysłowych, sejsmicznych itp.) oraz strukturalne modele reprezentacji wiedzy, oparte na lingwistycznej teorii zależności pojęciowej Rogera Schanka, wykorzystywane

do konstrukcji wielkich baz wiedzy w systemach przetwarzania języka naturalnego (np. w chatbotach i przeszukiwarkach internetowych).

Pięć powyżej omówionych podejść paradygmatycznych w sztucznej inteligencji zalicza się do jednego z dwóch jej podobszarów, mianowicie do symbolicznej sztucznej inteligencji (ang. *symbolic AI*), w której zakłada się, że: możliwa jest reprezentacja wiedzy w systemie inteligentnym *explicite*, wiedzę można reprezentować w sposób symboliczny, a symulację operacji umysłowo-poznawczych można przeprowadzić za pomocą formalnych operacji na zapisach symbolicznych. Drugim podobszarem sztucznej inteligencji jest inteligencja obliczeniowa (ang. *Computational Intelligence*, CI), w której zakłada się, że: wiedzę można reprezentować za pomocą informacji liczbowej, przetwarzanie wiedzy oparte jest głównie na obliczeniach (numerycznych), a wiedza w systemie może mieć charakter niejawny. Trzy omówione poniżej podejścia paradygmatyczne należą do tego właśnie podobszaru.

Najstarsze podejście inteligencji obliczeniowej stanowią modele inspirowane matematyką, gdyż początki rozpoznawania wzorców (nie należy mylić z wyżej wspomnianym syntaktycznym rozpoznawaniem wzorców), będącego największą (również w sensie praktycznych zastosowań) grupą metod w ramach tych modeli, można datować na lata 30. XX wieku – a zatem zanim pojawił się formalnie obszar badawczy sztucznej inteligencji. Wspólną cechą tych modeli jest wykorzystanie różnych działów matematyki (głównie: analizy matematycznej oraz rachunku prawdopodobieństwa) do rozwiązywania problemów: rozpoznawania obiektów/zdarzeń/procesów, automatycznego podziału ogromnych zbiorów obiektów na podzbiory reprezentujące określone kategorie (do każdego z takich podzbiorów należą jedynie obiekty podpadające definicyjnie pod odpowiadającą mu kategorię) czy też wnioskowania opartego na logikach wielowartościowych, rozmytych itp., schematach niemonotonicznych lub modelach statystyczno-podobnych, w sytuacji kiedy wiedza jest niepewna, niepełna lub też kategorie są zdefiniowane w niejednoznaczny sposób. Pewnym ograniczeniem tych modeli jest to, iż – w zasadzie – analizowane obiekty/zdarzenia/procesy mogą być charakteryzowane jedynie przez dane o charakterze liczbowym (a nie symbolicznym, kategoryalnym, jakościowym).

Najbardziej znanym (i zarazem najpowszechniej stosowanym w praktyce) reprezentantem drugiego podejścia inteligencji obliczeniowej –

modeli konekcyjnych – są sztuczne sieci neuronowe, za pomocą których próbuje się symulować działanie biologicznej sieci neuronowej, jaką jest mózg. Prekursorem tego podejścia był Frank Rosenblatt, który na początku lat 60. XX wieku skonstruował pierwszą sztuczną sieć neuronową, nazwaną Perceptronem. Szerokie praktyczne zastosowania tego podejścia powodują czasem wręcz utożsamianie obszaru sztucznych sieci neuronowych z obszarem sztucznej inteligencji, co oczywiście jest nieprawidłowe, gdyż pomimo ogromnych sukcesów w ich wykorzystaniu do rozwiązywania wielu problemów podejście to ma również swoje słabe strony (głównie: niejawnosć reprezentowanej wiedzy).

Trzecie podejście inteligencji obliczeniowej to obliczenia ewolucyjne, z których najbardziej znaną grupą metod są algorytmy genetyczne, rozwijające się dynamicznie po opublikowaniu przez Johna Hollanda monografii *Adaptation in Natural and Artificial Systems* w 1975 roku. Podstawowymi założeniami obliczeń ewolucyjnych są: możliwość konstrukcji przestrzeni potencjalnych rozwiązań danego problemu oraz możliwość symulacji procesów ewolucji gatunków, których poszczególne osobniki (lepiej lub gorzej przystosowane do „środowiska” tej przestrzeni) odpowiadają (lepszym lub gorszym) rozwiązaniom tego problemu. W celu konstrukcji systemów inteligentnych opartych na obliczeniach ewolucyjnych wykorzystuje się instrumentarium Darwinowskiej teorii ewolucji (populacje, genotypy osobników, fenotypy osobników, procesy selekcji osobników, krzyżowanie osobników, mutacja genotypu itp.) do implementacji odpowiednich operacji obliczeniowych (a także struktur danych). W ramach obliczeń ewolucyjnych, oprócz wspomnianych algorytmów genetycznych, rozwinęły się komplementarne do nich modele: strategii ewolucyjnych, które stosowane są w zadaniach optymalizacji numerycznej, programowania ewolucyjnego, w których dopuszcza się dowolny sposób reprezentacji genotypu osobnika (a nie jedynie: reprezentację sekwencyjną) oraz programowania genetycznego, którego celem jest nie tyle poszukiwanie rozwiązania instancji problemu, ile poszukiwanie metaschematu rozwiązywania (solvera) danego problemu (Flasiński, 2016). Do zagadnienia tego wrócimy jeszcze w następnym rozdziale.

Podsumowując, omówione powyżej podejścia paradygmatyczne w sztucznej inteligencji traktowanej jako obszar badań pokrywają w zasadzie cały jej obszar. Jak zatem widać, począwszy od lat 50.

ubiegłego wieku sztuczna inteligencja rozwijała się jako niejednorodny obszar badań z punktu widzenia wykorzystania różnych metodyk będących punktem wyjścia do konstrukcji systemów inteligentnych. Metodyki te oparte były (i są nadal) na różnych dyscyplinach nauki (psychologia (poznawcza), filozofia (umysłu), logika, lingwistyka, matematyka, neurofizjologia, biologia). W rzeczy samej, sztuczna inteligencja jako obszar badań jest, z metodologicznego punktu widzenia, „zlepkiem” silnie różniących się podobszarów badawczych, które łączy jedynie cel – konstrukcja systemów informatycznych (systemów robotyki), rozwiązujących problemy, jakie w przypadku człowieka wymagają inteligencji.

Ujęcie problemowe pojęcia

W niniejszej sekcji zidentyfikowane zostaną podstawowe zagadnienia i problemy, występujące w trakcie konstruowania (informatycznych) systemów sztucznej inteligencji zgodnie z podejściami paradygmatycznymi przedstawionymi w ujęciu historycznym w poprzednim rozdziale. Wcześniej jednak wprowadzimy fundamentalną zasadę metodologiczną konstrukcji systemów sztucznej inteligencji. Otóż systemy sztucznej inteligencji są systemami informatycznymi, w których dokonuje się symulacji (czyli: przybliżonego odtwarzania) inteligentnego działania (przy czym przez działanie rozumie się tutaj tak wykonywanie pewnych czynności, jak i – czysto teoretyczną – działalność umysłowo-poznawczą) pewnego obiektu (np. człowieka) (lub zespołu obiektów – co ma miejsce np. w przypadku obliczeń ewolucyjnych). Oznacza to, że konstruktorzy systemów sztucznej inteligencji projektują je jako informatyczne symulatory, opierając się na modelach inteligentnego działania, które to modele są „dostarczane” przez psychologów (poznawczych), filozofów (umysłu), logików, lingwistów, matematyków, neurofizjologów, biologów i przedstawicieli tych dyscyplin nauki, jakie mogą być pomocne w procesie konstrukcji systemów AI. Charakterystyki problemowej zagadnienia sztucznej inteligencji oraz identyfikacji podstawowych „problemów otwartych” w tym obszarze badań dokonamy zgodnie z wprowadzonym powyżej podziałem na dwa główne podobszary, tj. symboliczną sztuczną inteligencję oraz inteligencję obliczeniową w odniesieniu do każdej z grup metod z osobna.

Metody należące do podejścia symulacji kognitywnej, nazywane metodami przeszukiwania przestrzeni stanów, oparte są na następujących założeniach. Po pierwsze, dla problemu, którego rozwiązanie chcemy symulować, należy zdefiniować zasady konstruowania tzw. abstrakcyjnego modelu opisu (reprezentacji). Oznacza to, że dla każdej instancji tego problemu musimy potrafić wygenerować taki abstrakcyjny model. W przypadku pewnych problemów takie zadanie jest trywialne, np. w przypadku gier, problemów matematycznych itp., gdyż w tych przypadkach problemy zdefiniowane są właśnie w wysoce sformalizowany sposób. W przypadku problemów, które słabo poddają się formalizacji, mamy do czynienia z fundamentalnym (metodologicznym) ograniczeniem stosowalności tego podejścia. Z drugiej strony człowiek takie problemy rozwiązuje, gdyż zgodnie z psychologią poznawczą posiada zdolność tworzenia (abstrakcyjnych) reprezentacji umysłowych (struktur poznawczych), którym odpowiadają w systemach sztucznej inteligencji właśnie wymienione wyżej abstrakcyjne modele opisu. Zatem to rezultaty badań w obszarze psychologii poznawczej będą warunkować w najbliższej przyszłości rozwój metod symulacji kognitywnej w sztucznej inteligencji. Drugie poważne ograniczenie w rozwoju metod symulacji kognitywnej należy do obszaru informatyki. Otóż, kiedy możliwa jest reprezentacja problemu za pomocą abstrakcyjnego modelu opisu, na drugim etapie definiowania metody symulacji kognitywnej definiuje się tzw. przestrzeń stanów, która z matematycznego punktu widzenia jest grafem. Każdy węzeł grafu reprezentuje „stan” (krok) w procesie rozwiązywania problemu (wyróżnionymi stanami są: tzw. stany inicjalne odpowiadające sytuacjom, w których rozpoczyna się rozwiązywanie problemu – np. w grze w szachy jest jeden stan inicjalny i reprezentuje on inicjalną konfigurację na szachownicy – oraz stany końcowe zwykle odpowiadające sytuacjom, w których problem został rozwiązany). Każda krawędź grafu – łącząca dwa „stany”-węzły – odpowiada pojedynczej czynności w trakcie rozwiązywania problemu (w grze w szachy będzie to pojedynczy ruch). Oczywiście przestrzeń stanów jest konceptem teoretycznym, gdyż dla nietrywialnego problemu jest ona ogromnym grafem. Graf ten, a właściwie jego „szkielet” w postaci strukturalnej drzewa (tj. grafu acyklicznego), dana metoda symulacji kognitywnej generuje „cząstkowo” w trakcie rozwiązywania konkretnej instancji problemu. Ze względu na potencjalnie ogromne rozmiary takiego grafu podstawowym

problemem jest tutaj opracowanie takiej strategii generacji tego grafu, żeby był on możliwie najmniejszy. Zagadnienie to jest podstawowym drugim „problemem otwartym” w ramach symulacji kognitywnej.

Można stwierdzić, że w obszarze konstrukcji systemów służących do przeprowadzania wnioskowania na podstawie modeli logiki matematycznej rozwój efektywnych metod sztucznej inteligencji jest w wysokim stopniu satysfakcjonujący. Wynika to ze wspomnianego rozwoju logiki matematycznej w XX wieku, którego rezultaty są z wielkim sukcesem „konsumowane” przez obszar sztucznej inteligencji. Nie oznacza to, że w obszarze zastosowań logiki w sztucznej inteligencji nie są prowadzone ważne i interesujące badania (np. w zakresie logik: rozmytych, niemonotonicznych, temporalnych, zagadnień wnioskowania: przestrzennego, kauzalnego, „zdroworozsądkowego” (*common-sense reasoning*), itp.), ale wydaje się, że nie napotykają one na fundamentalne bariery rozwojowe, tak jak to dzieje się w innych podejściach do sztucznej inteligencji.

Podobna sytuacja, jak w przypadku metod opartych na logice matematycznej, ma miejsce w podobszarze modeli opartych na wiedzy regułowej. Systemy sztucznej inteligencji konstruowane na bazie tych modeli, nazywane eksperckimi systemami regułowymi (*expert rule-based systems*), służą do kodyfikacji wiedzy ograniczonej do dobrze określonego obszaru zastosowań, która to wiedza ma charakter wiedzy proceduralnej tak, jak jest ona zdefiniowana w psychologii poznawczej. Natomiast z punktu widzenia informatyki systemy takie działają w podobny sposób jak systemy oparte na logice matematycznej, gdyż zbiór reguł w takim systemie odpowiada zbiorowi formuł logicznych mających postać implikacji. Eksperckie systemy regułowe stanowią prawdopodobnie najszerszą klasę systemów sztucznej inteligencji wdrożonych do praktycznych zastosowań (w przemyśle, handlu, usługach, zarządzaniu itd.). Założenia teoretyczne konstrukcji takich systemów zostały opracowane w latach 70. i 80. ubiegłego wieku i nadal z powodzeniem są używane do konstruowania języków/środowisk programowania (a właściwie: języków/środowisk inżynierii wiedzy). Pewną zaletą tego podejścia w stosunku do podejścia opartego na logice matematycznej jest większa czytelność (rozumiałość) baz wiedzy opartych na regułach w porównaniu z bazami wiedzy zawierającymi formuły logiki matematycznej.

Podczas gdy modele oparte na wiedzy regułowej służą do reprezentacji i wnioskowania na bazie wiedzy proceduralnej, strukturalne

modele reprezentacji wiedzy są wykorzystywane głównie do reprezentacji wiedzy deklaratywnej tak, jak jest ona zdefiniowana w psychologii poznawczej. Również i w przypadku tego podejścia, którego najbardziej dynamiczny rozwój przypada na lata 70. i 80. ubiegłego wieku, podstawy teoretyczne opracowane we współpracy z psychologami, językoznawcami oraz psycholingwistami pozwalają do dzisiaj konstruować duże bazy wiedzy wykorzystywane w wielu praktycznych zastosowaniach. Do najbardziej znanych modeli w tym obszarze należy zaliczyć: sieci semantyczne Allana M. Collinsa i Rosa Quilliana, za pomocą których możemy konstruować tzw. ontologie (formalne reprezentacje wiedzy dziedzinowej oparte głównie na relacjach pomiędzy pojęciami/kategoriami), systemy ramowe (ramy) (strukturalne reprezentacje wiedzy bazujące na stereotypowych schematach) Marvinina Minsky'ego, a także skrypty (strukturalne reprezentacje wiedzy – tym razem: proceduralnej – oparte na stereotypowych sekwencjach zdarzeń) Rogera Schanka i Roberta P. Abelsona. Również w przypadku strukturalnych modeli reprezentacji wiedzy możemy mówić o „domknięciu” (w sensie możliwości dalszego rozwoju) tego obszaru badawczego.

W przypadku syntaktycznej analizy wzorców opartej na koncepcji gramatyk generatywnych Noama Chomsky'ego, pomimo szerokiego spektrum ich zastosowań wymienionych w poprzednim rozdziale, nadal nie udało się w sposób satysfakcjonujący rozwiązać fundamentalnego problemu wnioskowania gramatycznego (pomimo opracowania szeregu efektywnych metod dla pewnych podklas problemów). Zagadnienie to jest typowym reprezentantem trudnych problemów sztucznej inteligencji i może być zdefiniowane jako próba konstrukcji algorytmu, próbującego na podstawie próbki zdań (albo innych wzorców formalnych) należących do niezdefiniowanego formalnie języka odtworzyć gramatykę generatywną (a zatem formalizm definiujący ten nieznaną język), która: w prostszej wersji tego problemu – generuje tę próbkę, względnie – w trudniejszej wersji tego problemu – generuje wszystkie (domnimane) zdania tego języka. Możliwość konstrukcji takich algorytmów (dla rozmaitych klas gramatyk generatywnych) stanowi aktualnie najpoważniejszą barierę fundamentalną dla dalszego rozwoju metod należących do tego podejścia w sztucznej inteligencji (Flasiński, 2019).

Podstawowe modele i metody najstarszego podejścia w ramach podobszaru inteligencji obliczeniowej, tj. rozpoznawania wzorców (zwanego

też: rozpoznawaniem obrazów) zostały opracowane w latach 60. i 70. ubiegłego wieku. Takie metody jak: liniowa analiza dyskryminacyjna, liniowy dyskryminator Fishera, klasyfikator Bayesa, metoda k-NN, metoda wektorów wspierających, klasyfikator drzew decyzyjnych czy też algorytm klasteryzacji k-średnich, stanowią dzisiaj nie tylko standardowe modele sztucznej inteligencji, ale traktowane są jako „klasyka” informatyki. Chociaż również w przypadku tego podejścia można mówić o „domknięciu” obszaru badawczego, to cały czas ciekawe wyniki teoretyczne i aplikacyjne uzyskiwane są w badaniach znajdujących się na pograniczu zagadnień rozpoznawania wzorców, uczenia maszynowego oraz sztucznych sieci neuronowych.

W przypadku obszaru obliczeń ewolucyjnych dwie wspomniane powyżej grupy metod, a mianowicie algorytmy genetyczne oraz strategie ewolucyjne, należy zaliczyć do dobrze ugruntowanych teorii, stanowiących standardowe modele w tym obszarze. Natomiast w przypadku pozostałych dwóch grup metod, tj. programowania ewolucyjnego oraz programowania genetycznego sytuacja jest trudniejsza, gdyż celem konstrukcji metod sztucznej inteligencji w ramach tych dwóch grup jest nie tyle opracowanie – w przypadku trudnego problemu – metody, umożliwiający znalezienie satysfakcjonującego (zgodnie z przyjętymi kryteriami) rozwiązania dla zadanej instancji problemu, ale raczej opracowanie metamedody, która generowałaby „rozwiązywacz” (*solver*) problemów należących do pewnej klasy. Potencjalne zastosowania tak ambitnie sformułowanego zagadnienia obejmowałyby system sztucznej inteligencji, który pisze programy komputerowe, projektuje układy elektroniczne czy też „odkrywa” nowe farmaceutyki. W tym obszarze badań należy spodziewać się jeszcze wielu przełomowych rezultatów.

Najbardziej znaną grupą modeli inteligencji obliczeniowej są sztuczne sieci neuronowe. W przypadku tej grupy w ostatniej dekadzie odnotowano najszybszy rozwój. Co więcej, dotyczy on fundamentalnej do tej pory bariery metodologicznej dla sztucznych sieci neuronowych, a mianowicie możliwości przetwarzania przez sieci informacji o charakterze symbolicznym. Sztuczny neuron, z istoty swojej – jako model neuronu biologicznego – służy do przetwarzania sygnałów (ciągłych), a więc informacji o charakterze numerycznym (liczbowym). Przez pierwsze pół wieku wykorzystywania sztucznych sieci neuronowych do rozwiązywania praktycznych problemów w sztucznej inteligencji była to

podstawowa przeszkoda, uniemożliwiająca wykorzystanie ogromnego potencjału sieci neuronowych związanego ze zdolnością zapamiętywania przez nie ogromnych zbiorów danych. Chociaż pierwsze próby zastosowania sieci neuronowych w przetwarzaniu języka naturalnego (a więc informacji o charakterze właśnie symbolicznym) można datować na lata 90. ubiegłego wieku (sieci rekurencyjne Jeffreya L. Elmana), to dopiero opracowanie w 2017 roku koncepcji tzw. transformera, opartej na zaawansowanej architekturze obliczeniowej z wykorzystaniem modułów wielowarstwowych sieci neuronowych, umożliwiło pokonanie tej fundamentalnej bariery metodologicznej. W efekcie zaowocowało to możliwością konstrukcji tzw. dużych modeli językowych (ang. *Large Language Models*, LLM), np. typu GPT (ang. *Generative Pre-trained Transformer*) czy też chatbotu ChatGPT (Yurafsky & Martin, 2023). W najbliższej przyszłości należy spodziewać się dynamicznego rozwoju sztucznej inteligencji właśnie w tym jej obszarze.

Refleksja systematyczna z wnioskami i rekomendacjami

Sztuczna inteligencja jest najbardziej dynamicznie rozwijającym się obszarem badań w dyscyplinie informatyka. Dotyczy to każdej z wymienionych w poprzedniej sekcji grup metod (choć śledząc historię rozwoju sztucznej inteligencji, można zauważyć okresy, w których pewne grupy metod wyraźnie znajdowały się w stanie regresu, podczas gdy inne rozwijały się w bardzo szybkim tempie). Pojawia się tutaj fundamentalne pytanie, które ważne jest nie tylko ze względu na możliwości praktycznych zastosowań systemów inteligentnych (informatyka, robotyka), ale także na wydźwięk i znaczenie możliwej odpowiedzi dla filozofii, psychologii czy też kognitywistyki, a mianowicie: czy sztuczna inteligencja może dorównać (względnie przewyżżyć) inteligencji „naturalnej”, czyli inteligencji człowieka? (Inteligencja człowieka znajduje się na szczycie inteligencji znanych nam empirycznie, czyli w świecie przyrody. Oczywiście, filozofowie należący do pewnych szkół mogą jeszcze rozważać inteligencje stojące wyżej, np. substancje oddzielone). Naukowcy działający w obszarze sztucznej inteligencji są podzieleni na dwie grupy, jeśli chodzi o odpowiedź na to pytanie. Przedstawiciele

silnego podejścia do sztucznej inteligencji (ang. *strong AI*) twierdzą, że nie istnieją przeszkody, żeby skonstruować komputerowy system inteligentny, który będzie w stanie działać tak, jak ludzki mózg wraz z jego wszystkimi czynnościami umysłowymi oraz poznawczymi (włącznie z fenomenami psychicznymi, np. świadomością, motywacją, uważnością, odczuciami itp.). Z kolei zwolennicy słabego podejścia do sztucznej inteligencji (ang. *weak AI*) są zdania, że chociaż komputerowy system inteligentny potrafi wykonywać pewne symulacje czynności umysłowo-poznawczych i są one przydatne w testowaniu hipotez dotyczących funkcjonowania ludzkich procesów percepcyjnych, poznawczych i umysłowych, to jednak nie będziemy w stanie odtworzyć w systemie informatycznym (robotyki) w pełni tych procesów. (Przegląd różnych poglądów w obszarze filozofii umysłu należących do *strong AI* oraz *weak AI* można znaleźć w: Flasiński, 2016).

W kontekście powyższego pytania o możliwość sztucznej inteligencji – jako pełnej realizacji inteligencji ludzkiej – należy zastanowić się, czy istnieją jakieś fundamentalne potencjalne bariery, które uniemożliwiałyby taką symulację. Zanim jednak spróbujemy odpowiedzieć na tak postawione pytanie, należałoby się zastanowić nad samym pojęciem inteligencji. U Flasińskiego (2016) dokonano syntezy różnych definicji inteligencji spotykanych w literaturze z obszaru psychologii poznawczej, co pozwoliło określić inteligencję jako zespół zdolności umożliwiających:

- adaptację do zmieniającego się otoczenia,
- aktywność poznawczą wyrażającą się w konstruowaniu struktur abstrakcyjnych (odpowiadających w zasadzie strukturom poznawczym, reprezentacjom umysłowym tak, jak są one ujmowane w psychologii poznawczej) i operowaniu nimi.

(Przegląd i charakterystyka takich zdolności na podstawie literatury z psychologii poznawczej znajduje się w: Flasiński, 2016).

Potencjalnych barier w rozwoju sztucznej inteligencji należy, oczywiście, szukać w drugiej składowej powyższej definicji inteligencji (pierwszy komponent definicji wydaje się spełniony przez współczesne systemy inteligentne w dużym stopniu i nie widać tutaj fundamentalnych barier, które mogłyby ograniczać dalszy rozwój sztucznej inteligencji w tym aspekcie).

Identyfikacja i charakterystyka takich barier została przeprowadzona i przedstawiona np. u Flasińskiego (2016). Wykorzystana została w tym

celu klasyfikacja podstawowych operacji poznawczych zgodnie z epistemologicznymi poglądami Arystotelesa i św. Tomasza z Akwinu. Otóż, zgodnie z tymi poglądami, można wyróżnić trzy generyczne operacje umysłowo-poznawcze (łac. *tres operationes rationis*): tworzenie pojęć (łac. *indivisibilium intelligentia*), wydawanie sądów (łac. *componere et dividere*) oraz przeprowadzanie rozumowań (łac. *rationare*) oznaczające przechodzenie od jednego sądu do drugiego zgodnie z regułami logiki. Spośród tych trzech generycznych operacji jedynie w przypadku przeprowadzania rozumowań można stwierdzić ich satysfakcjonującą realizację w systemach sztucznej inteligencji. Wynika to ze znakomitego rozwoju logiki w ubiegłym wieku. Rezultaty tego rozwoju dotyczą wszystkich podstawowych zagadnień, które są istotne dla konstrukcji efektywnych algorytmów symulujących różne aspekty rozumowań człowieka. Po pierwsze, algorytmizacja taka była możliwa dzięki osiągnięciu bardzo wysokiego stopnia formalizacji rachunków logicznych, wynikającego z wielkiej dbałości o precyzyjną definicję składni (syntaksy) tych rachunków (na co niewątpliwie wpłynęło pojawienie się nurtu filozofii analitycznej z jego silnym położeniem akcentu na zagadnienia językowe). Po drugie, modele logiki (matematycznej) były wprowadzane na podstawie paradygmatu obliczania symbolicznego, co stanowiło później duże ułatwienie dla informatyków konstruujących narzędzia informatyczne do konstrukcji systemów sztucznej inteligencji (np. język LISP oparty na rachunku lambda Alonza Churcha czy też język PROLOG oparty na metodzie wnioskowania rezolucyjnego w logice pierwszego rzędu). Po trzecie, bardzo ważny – z punktu widzenia możliwości zastosowania rachunków logiki do rozwiązywania problemów związanych z otaczającą nas rzeczywistością fizyczną – był fakt rozwoju semantyki logicznej/teorii modeli za sprawą wybitnych logików XX wieku, takich jak: Alfred Tarski, Thoralf Skolem czy też Kurt Gödel.

Znacznie gorzej sytuacja wygląda w przypadku drugiej generycznej operacji umysłowo-poznawczej, tj. wydawania sądów. Analizy tego zagadnienia dokonamy, podobnie jak Flasiński (2016), na podstawie Kantowskiej dystynkcji sądów. W przypadku sądów analitycznych *a priori* ich generacja jest dokonywana w systemach sztucznej inteligencji poprzez odniesienie się do wiedzy zgromadzonej w tzw. bazach wiedzy, będących w systemach sztucznej inteligencji odpowiednikiem baz danych w (standardowych) systemach informatycznych. Bazy

wiedzy są konstruowane za pomocą różnych podejść sztucznej inteligencji (omówionych w poprzednich rozdziałach). Przykładowo mogą one mieć charakter strukturalnych reprezentacji wiedzy (np. sieci semantyczne, systemy ramowe, skrypty), zbiorów formuł logiki pierwszego rzędu (lub innych modeli logiki matematycznej) itp. Generacja sądów analitycznych *a priori* w systemie sztucznej inteligencji wykorzystuje zatem (odtwórczo) istniejącą już w systemie (zgromadzoną wcześniej) wiedzę (co jest zgodne z sensem pojęcia sądu analitycznego *a priori* w ujęciu kantowskim). Sądy syntetyczne *a posteriori* są wyprowadzane na podstawie wcześniejszych doświadczeń. W systemach sztucznej inteligencji generacji takich sądów dokonuje się poprzez symulację logicznego schematu wnioskowania (reguły inferencji) indukcyjnego. Takiego schematu używamy, przykładowo, w trakcie uczenia sztucznych sieci neuronowych, uczenia klasyfikatorów w obszarze rozpoznawania wzorców (obrazów) czy też wnioskowania gramatycznego w syntaktycznym rozpoznawaniu wzorców. Zatem – również w przypadku sądów tego rodzaju – potrafimy symulować ich generację w systemach sztucznej inteligencji. Natomiast problem – w postaci bariery fundamentalnej – pojawia się w przypadku sądów syntetycznych *a priori*. Wynika to z faktu, że ani psychologia poznawcza ani filozofia umysłu (czy też ogólnie: epistemologia) nie potrafią odpowiedzieć na pytanie, jaki jest mechanizm powstawania takich sądów. (Zwykle fenomen powstawania sądu syntetycznego *a priori* określa się za pomocą tak słabo algorytmizowalnych pojęć, jak wgląd (w istotę rzeczy), intuicja intelektualna, czysta naoczność itp.).

Najgorzej sytuacja wygląda w przypadku operacji tworzenia pojęć (pojęciowania), rozumianej jako proces abstrahowania. W tym przypadku również ani psychologia poznawcza, ani filozofia umysłu nie potrafią opisać/wyjaśnić mechanizmu pojęciowania. Pewne nadzieje w obszarze tego zagadnienia były wiązane w latach 80. i 90. z ciekawymi koncepcjami „kategoryzacji ucieleśnionej” wprowadzonymi w lingwistyce kognitywnej George’a Lakoffa i Ronalda Langackera. Choć wydaje się, że koncepcje lingwistyki kognitywnej mogłyby stanowić obiecujący punkt wyjścia dla konstrukcji algorytmów symulacji ludzkiego pojęciowania, to niestety nadal – po niemal pół wieku od ich pojawienia się – nie udało się wprowadzić ich do obszaru konstrukcji systemów sztucznej inteligencji.

Podsumowując ten wątek rozważań, należy stwierdzić, że powyższe dwie fundamentalne bariery, związane z tworzeniem pojęć oraz wydawaniem sądów syntetycznych *a priori*, stanowią (przynajmniej na tym etapie rozwoju badań nad sztuczną inteligencją) przeszkodę w realizacji postulatu silnej sztucznej inteligencji.

Kończąc refleksję systematyczną na temat sztucznej inteligencji, należy wspomnieć jeszcze o dwóch zagadnieniach. Pierwsze zagadnienie dotyczy przeciwstawienia tzw. ogólnej sztucznej inteligencji (ang. *artificial general intelligence*, AGI) vs sztucznej inteligencji specjalizowanej (ang. *narrow AI*). Podczas gdy systemy sztucznej inteligencji należące do tej drugiej kategorii są konstruowane w wąsko pojętych obszarach symulacji pewnych zdolności umysłowo-poznawczych człowieka (np. przetwarzanie języka naturalnego, rozpoznawanie obrazów, systemy eksperckie diagnostyki medycznej), przez systemy AGI należałoby rozumieć systemy, które charakteryzują się inteligencją ogólną, tzn. zdolne do rozwiązywania jakichkolwiek zadań intelektualno-poznawczych, które człowiek może rozwiązywać. Należy zaznaczyć, że w przypadku (hipotetycznych) systemów AGI nie wymaga się od nich zwykle przejawiania innych fenomenów psychicznych, o których mówi się w definicji silnej AI (np. świadomości, uważności itp.). Drugie zagadnienie jest związane z pojęciem superinteligencji (ang. *superintelligence*), tzn. sztucznej inteligencji przewyższającej inteligencję ludzką. (Hipotetyczne) sztuczne superinteligentne systemy wiąże się zwykle z koncepcjami: osobliwości technologicznej Raya Kurzweila oraz transhumanizmu Nicka Bostroma. Wydaje się jednak, że to zagadnienie należy do obszaru futurologii.

BIBLIOGRAFIA

- Flasiński, M. (2016). *Introduction to Artificial Intelligence*. Cham: Springer.
- Flasiński, M. (2019). *Syntactic Pattern Recognition*. New Jersey–London–Singapore: World Scientific.
- Haugeland, J. (red.) (1985). *Artificial Intelligence: The Very Idea*. Cambridge: MIT Press.
- Nilsson, N.J. (1998). *Artificial Intelligence. A New Synthesis*. San Francisco: Morgan Kaufmann.

- Rich, E., & Knight, K. (1991). *Artificial Intelligence*. New York: McGraw-Hill.
- Russell, S., & Norvig, P. (2020). *Artificial Intelligence. A Modern Approach*. Boston: Pearson.
- Winston, P.H. (1992). *Artificial Intelligence*. Reading, MA: Addison-Wesley.
- Yurafsky, D., & Martin, J.H. (2023). *Speech and Language Processing*. Upper Saddle River: Prentice-Hall.

Dariusz Adamek
Uniwersytet Jagielloński w Krakowie
<https://orcid.org/0000-0002-5570-2259>
<https://doi.org/10.35765/slowniki.481>

Reakcje emocjonalne

Streszczenie

DEFINICJA POJĘCIA: Reakcje emocjonalne to forma złożonych i skoordynowanych zjawisk motorycznych, wspomagających je zmian czynnościowych w całym organizmie i zmian w sposobie działania samego układu nerwowego, w tym w szczególności tzw. układu wegetatywnego, koniecznych dla uniknięcia ważnych dla przeżycia zagrożeń lub związanych z realizacją pożądaných dla osobnika życiowo-ważnych celów zarówno w sposób fizyczno-realny, jak i mentalny/konceptualny.

ANALIZA HISTORYCZNA POJĘCIA: Terminy „reakcje emocjonalne”, „mózg emocjonalny” pojawiają się w literaturze od dawna, ale osobiste „spotkanie” autora niniejszego tekstu z tymi terminami ma dokładną datę – rok 2004, czyli moment ukazania się przeglądowego artykułu T. Dalgleisha pt. „The emotional brain” (Dalgleish, 2004).

UJĘCIE PROBLEMOWE POJĘCIA: Emocje i reakcje emocjonalne to zagadnienia złożone i wieloznaczne, wymagające badań interdyscyplinarnych (od neurologii poprzez psychologię do zoologii). W tym słownikowym haśle Autor odnosi się głównie do wybranych neurobiologicznych aspektów emocji i reakcji emocjonalnych. Przejawy emocji można dostrzec nawet u tak „prymitywnych” zwierząt jak bezkręgowce czy jeszcze „prymitywniejszych” jak jednokomórkowce (kluczowa jest tu kwestia definicji).

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Od czasu odkrycia przez Paula Broca stosunkowo niewielkiego obszaru w płacie czołowym lewej półkuli mózgu, którego uszkodzenie powoduje

zaburzenia mowy (nazywane afazją), datuje się przekonanie, że istnieją szczególne miejsca w mózgu, tzw. ośrodki, które wykonują specjalne zadania-funkcje, w tym także ośrodki odpowiedzialne za emocje. Otworzyło to drogę do ukształtowania się osobnego działu medycyny, jakim jest neurologia.

Słowa kluczowe: rodzaje emocji, motywacje, trójjedyny mózg, obwód Papeza, badania interdyscyplinarne

Definicja pojęcia

Reakcje emocjonalne to forma złożonych i skoordynowanych zjawisk motorycznych, wspomagających je zmian czynnościowych w całym organizmie i zmian w sposobie działania samego układu nerwowego, w tym w szczególności tzw. układu wegetatywnego, które:

- A. Po pierwsze umożliwiają, ułatwiają i usprawniają efektywną reakcję osobniczą na życiowo-ważne potrzeby lub wyzwania: uniknięcie zagrożenia-niebezpieczeństwa, przeżycie w sytuacji zagrożenia (takiej jak pojawienie się napastnika czy rywala do partnera seksualnego lub do pokarmu), zdobycie pożywienia, prokreacja, zaspokojenie pragnienia (wody), zaopiekowanie się potomstwem.
- B. Po drugie (to nieco bardziej dyskusyjna teza) umożliwiają/ułatwiają adaptację organizmu do zmian w otoczeniu lub w nim samym w sposób minimalizujący straty/szkody oraz optymalizujący szanse na sukces w przyszłości.
- C. Po trzecie (to również raczej dyskusyjna teza) są istotą „motoru motywacyjnego”, innymi słowy są na bardzo podstawowym poziomie konieczne, choć oczywiście niewystarczające, aby „nam się chciało chcieć”.

Panuje niemały zamęt, gdy chodzi o aparat pojęciowy związany z emocjami i tym, co uważamy za ich przejawy. Spór dotyczy samych definicji rodzajów emocji i ustalenia, które z nich są bardziej pierwotne/proste, a które raczej wtórne/złożone. Na przykład można zapytać: czym są afekty, popędy? Które emocje uznajemy za podstawowe? Dyskutowane są takie pojęcia/terminy jak „afekt rdzenny”, „inteligencja emocjonalna”. Według Jaaka Pankseppa mózgi ssaków wykazują obecność systemów emocjonalnych, z których cztery, a mianowicie: poszukiwanie (*seeking*), strach (*fear*), wściekłość (*rage*), panika (*panic*), są obecne od urodzenia, a nieco później pojawiają się: żądza (*lust*), opiekuńczość (*care*), zabawa-wesołość (*play*). Zauważmy, że nie wymieniono tu np. afektów takich jak: uczucia nienawiści, zazdrości, miłości i wiele innych. Kwestia ta została rozwinięta w książce pod redakcją Davisa i Pankseppa (2017).

Sądzę, że daje się obronić teza, że analogicznie jak to jest np. w przypadku świadomości, którą najbardziej zdroworozsądkowo jest badać poprzez interdyscyplinarną obserwację tzw. korelatów świadomości, tak

w przypadku emocji optymalnym rozwiązaniem wydaje się analizowanie ich „przejawów” – czyli reakcji emocjonalnych. Podobnie bowiem jak to jest w przypadku zagadnienia świadomości, do którego odnosi się słynny i nierozstrzygalny dylemat T. Nagela „Jak to jest być nietoperzem”, szeroko analizowany przez J. Bremera (Bremer, 2005), można go sparafrazować pytaniem: „Jak to jest być smutnym nietoperzem?” (albo nietoperzem zadowolonym, rozwścieczonym itp.). Emocje są niewątpliwie komunikowalne, ale co one rzeczywiście oznaczają dla indywidualnego człowieka czy zwierzęcia?

Jeśli, jak wspomniałem, zastosuje się ewolucjonistyczne podejście do reakcji/przejawów emocjonalnych rozważanych w bezpośrednim kontekście biologicznym – w tym neurobiologicznym – prawdopodobnie będzie można się zgodzić z następującą tezą/definicją, którą zaproponowałem na wstępie (proszę zwrócić uwagę, że zamiast terminu „emocje” najczęściej używam określenia „reakcje emocjonalne” gdyż można je obserwować i tylko poprzez nie można wnosić o emocjach człowieka czy innego zwierzęcia, ale też nie przeczę, że w dalszym tekście „emocje” i „reakcje emocjonalne” są często traktowane jako zamienniki): Odwołując się do podanej na wstępie sugerowanej definicji pojęcia „reakcji emocjonalnych”, rozważmy następujące aspekty definicji wymienione w podpunktach (A, B, C).

O ile przykłady ilustrujące pkt A są raczej mniej lub bardziej oczywiste, o tyle ilustracją pkt. B może być stan/emocja smutku lub lęku. Takie stany emocjonalne mogą ułatwiać przeżycie w sytuacji stresu lub po silnym stresie, w którym osobnik nie ma możliwości aktywnego przeciwdziałania negatywnym czynnikom czy działaniom, próbuje jednak przynajmniej albo zminimalizować szanse na zwrócenie na siebie uwagi np. przez bezruch (u zwierząt dobrze znana jest reakcja „zamrożenia” – *freezing* – przyjmowana zazwyczaj jako bardziej optymalny wariant zachowania w celu uniknięcia ataku przez napastnika), albo stara się np. przez manifestację smutku sprowokować postawy uczuciowe w rodzaju współczucia (empatii), albo jeszcze dla siebie korzystniej – oczekiwać reakcji pomocy od innych osobników w określonej grupie, być może z ewentualnym odnowieniem szans na późniejszy sukces (np. reprodukcyjny).

Rola motywacyjna emocji (pkt C powyżej) – a ściślej: tych pozytywnych – nie jest oczywista. Nie jest kontrowersyjne przeświadczenie, że to emocje popychają nas do wspomnianych wyżej powrotów

do miejsc dzieciństwa, do wspólnego świętowania Bożego Narodzenia itp. Ale czy przymus wewnętrzny, jaki pcha łososie do Morza Sargasowego albo ptaki do wędrówki do miejsc lęgowych, a nawet miliony niektórych gatunków motyli do rocznych „podróży”, których dystanse liczone są w tysiącach kilometrów, można też zakwalifikować jako emocjonalnie motywowane? Wprowadzam kolejne rozróżnienie, mówiąc, że powszechna opinia skłonna jest uznawać je raczej za napędzane „wrodzonym” (?) instynktem prokreacji czy instynktem przeżycia (a więc warunkowanym zdobyciem wody i pożywienia). Niemniej pojawią się pytania: (a) Czy jest to jednak zupełnie inna motywacja niż „pogoń” człowieka za tym, aby się najeść, napić i dokonać „wiadomego aktu”, zdobyć wyższy status społeczny itp.? (b) Czy dążąc po raz pierwszy do wykonania „wiadomego aktu”, człowiek różni się bardzo od łososia, który dąży do szczęścia złożenia ikry poprzez akt, o którym nie ma pojęcia w sensie własnego, osobistego doświadczenia? Zgoda, diametralnie inna jest sytuacja, gdy „popędowe”, intuicyjne dążenie do spełnienia aktu (np. prokreacji) u danego osobnika ma szansę na realizację tylko jeden jedyny raz w życiu osobniczym, tak jak to jest u łososia, ale już w sytuacji, gdy intuicyjnie-popędowo warunkowany akt ma szansę się powtórzyć, całkowicie zmienia się kontekst. Nawet relatywnie prymitywne zwierzę uzyskuje dostęp do osobistego doświadczenia, które zachęca lub odstręcza. W ten sposób powstają (pozytywne lub negatywne) mechanizmy motywacji, których dotyczy także możliwość amplifikacji lub tłumienia. Konieczne jest tu spełnienie drugiego (poza możliwością powtórzeń) warunku, a mianowicie – pamięci. Oczywiście pamięci nabytej, co różni ją od innych form pamięci, w szczególności pamięci polegającej na zakodowanych wrodzonych „algorytmach”. Takim algorytmem wrodzonym jest przykładowo pierwotna motywacja „nakazująca” noworodkowi ssaka poszukiwanie sutków. W kolejnych epizodach „poszukiwania” sutka i karmienia mechanizmy pamięci amplifikują pierwotny odruch i ta możliwość zależnej od pamięci amplifikacji danego aktu (w tym przypadku karmienia piersią) staje się emocją (reakcją emocjonalną) w znaczeniu pkt. C. Chodzi więc tu o emocję motywującą, „napędową”. Niemniej jednak należy odróżniać wspomniane motywacje od motywacji prowadzących do świadomych ludzkich działań, zwłaszcza tych długofalowych, opartych na wielorakich wspomnieniach i nakierowanych na określone cele.

Analiza historyczna pojęcia

Odpowiadając na powyższe pytania (a, b) o różnicę pomiędzy ludzkimi reakcjami emocjonalnymi a zwierzęcymi, pragnę zauważyć, że dokładnie nie wiadomo, od kiedy historycznie łączono tę motywacyjną rolę emocji z sercem (rozumianym metaforycznie, choć początkowo raczej dosłownie). Z pewnością nastąpiło to w czasach antycznych. Wręcz utożsamiano emocje/emocjonalność z tym narządem. Literatura i sztuka dostarczają niezliczonych przykładów gdzie serce symbolizuje emocje, w tym zwłaszcza te motywacyjne, napędowe i raczej przeciwstawiane rozumowi czy umysłowi (z reguły postrzeganemu jako „chłodny” w przeciwieństwie do „gorącego” serca). Mój ulubiony cytat wskazujący na taką motywacyjną rolę serca (a właściwie emocji) to początek *Wyznań* św. Augustyna, w którym znajdujemy stwierdzenie: „Uczyniłeś nas dla siebie, o Panie, i niespokojne jest serce człowieka, dopóki nie spocznie w Tobie” (Augustyn, 1987). Zawarta w nim ukryta teza brzmi, że to Bóg zamyślił emocje jako mechanizm motywacyjny w „projekcie” człowieka. Przedstawiona wyżej próba definicji reakcji emocjonalnych scala w szczególności to, co nazywa się „popędami” oraz „emocjami” (Ciccharelli & White, 2023)

Od odkrycia przez Paula Broca (1824–1980) stosunkowo niewielkiego obszaru w płacie czołowym lewej półkuli mózgu, którego uszkodzenie powoduje zaburzenia mowy (nazywane afazją), datuje się przekonanie, że istnieją szczególne miejsca (ośrodki) w mózgu, które wykonują specjalne zadania-funkcje (LaPointe, 2012). Otworzyło to drogę do ukształtowania się osobnego działu medycyny, jaką jest neurologia, która dzisiaj wchodzi w skład interdyscyplinarnie rozumianych neuronauk. Od dawna też interesowało lekarzy i naukowców, czy i które struktury mózgu odpowiedzialne są za emocje, zwłaszcza te najsilniejsze (jak uczucie miłości czy nienawiści). Jakie są „emocjonalne uzwojenia” mózgu? Dominowało podejrzenie/przekonanie, że nie da się wskazać „anatomicznego ośrodka” jakiejś konkretnej emocji, jak w przypadku ośrodka Broca. Może zatem zadanie takie mogłaby spełniać nieco rozproszona „konstrukcja” anatomicznie i (co za tym można wnioskować) funkcjonalnie połączonych struktur mózgu? Pierwszą taką neuroanatomiczną koncepcję emocji, którą zaproponowano, był, jak go dzisiaj nazywamy (od nazwiska autora, Jamesa Papeza (1883–1958),

amerykańskiego neuroanatomia), tzw. obwód Papeza (Papez, 1937). Tworzą go następujące struktury: hipokamp – sklepienie (*fornix*) – ciała suteczkowate – droga suteczkowo-wzgórzowa (*tractus mammillothalamicus*) – jądro przednie wzgórza – włókna projekcyjne ze wzgórza do zakrętu obręczy (*gyrus cinguli*) – zakręt obręczy – (włókna projekcyjne z zakrętu obręczy do hipokampa – hipokamp (co „domyka” ten krąg połączeń).

W roku 2004 ukazał się przeglądowy artykuł T. Dalgleisha pt. „The emotional brain” (Dalgleish, 2004). Artykuł ten oraz jeszcze dwie publikacje książkowe (Purves, 2008; Bear, 2016) dostarczają większość historycznie znanych podstaw konceptualnych, ale także znaczną część inspiracji do własnych przemyśleń, których owoce przedstawiam w poniższym tekście. W tym zakresie, modelowe jest porównanie klasycznych teorii emocji.

Pierwsza z nich, zwana teorią Jamesa-Langego (twórcy: William James, Carl Lange –koniec XIX w.) zakłada, że objawy generowane poprzez działanie głównie autonomicznego układu nerwowego odczuwamy jako emocję. Druga, teoria Cannona-Barda (Walter Cannon, Philip Bard – lata 20. XX w.), zakłada z kolei, że czynnik postrzegany jako „emocjonalny” wywołuje reakcję emocjonalną, czyli reakcję mającą swoje źródło w centralnym układzie nerwowym prowadzącą do odpowiednich zmian w funkcji różnych narządów i reakcji ruchowych – somatycznych (np. ucieczka) (Cannon, 1927).

Mówiąc obrazowo (może brzmieć to nieco trywializująco), wg teorii Jamesa-Langego, jeśli ktoś biegnie, a serce mu silniej bije, poziom adrenaliny i metabolizm gwałtownie rosną, oznacza to najprawdopodobniej, że jego mózg „odczuwa” neuronalny korelat strachu. Według teorii Cannona-Barda porządek jest niejako odwrócony: ktoś zobaczył groźne zwierzę, jego mózg „zinterpretował” egzystencjalne niebezpieczeństwo, jakie wynika z tego faktu, czyli uznał, że należy się bać, i uruchomił układ wegetatywny oraz szereg innych mechanizmów koniecznych/przydatnych do tego, aby pokonać tę sytuację (w zależności od oceny, aktywnie lub biernie, czyli przez walkę lub ucieczkę). Na pierwszy rzut oka pierwsza teoria wydaje się „dziwna”, jednak dla zewnętrznego obserwatora jej implementacja np. w badaniu naukowym zwierząt ludzkich czy nieludzkich jest bardziej konstruktywna. Nie wiemy, co naprawdę czuje małpa, gdy postrzega węża (klasyczny model „strachu”), ale jeśli gwałtownie

ucieka (przemieszcza się i, jeśli jest drzewo, to z reguły czyni to w kierunku pionowym...), to mamy prawo zakładać, że czuje strach. To sytuacja analogiczna jak ta dotycząca świadomości, w przypadku której tak naprawdę nie wiemy, czym i kiedy jest. Dysponujemy jedynie założeniami, że pewne „korelaty” implikują jej istnienie czy raczej „posiadanie”.

Ujęcie problemowe pojęcia

Uwagi wstępne

Od 20 lat w ramach fakultatywnego kursu podstaw neurobiologii dla studentów medycyny poświęcam jeden interdyscyplinarny wykład zagadnieniu „mózgowych” czy raczej „nerwowych” mechanizmów emocji. W semestrze zimowym wykład ten celowo umieszczam w tygodniu bezpośrednio poprzedzającym święta Bożego Narodzenia, argumentując to przeświadczeniem, że są to najbardziej emocjonalne święta w naszym kręgu kulturowym, szczególnie w Polsce... Za jeden z argumentów „za”, jakie przedstawiam, służy *Scherzo h-moll* op. 20 F. Chopina z motywem kolędy *Lulajże Jezuniu*. Jest to przetworzony artystycznie symboliczny, mentalny „powrót” kompozytora do Bożego Narodzenia czasów dzieciństwa. Ta w sposób genialny zacytowana przez Chopina kolęda, wyłaniająca się pośród jakiegoś „okropieństwa” melodycznego, wściekłego zgiełku wydarzeń i nastrojów nieomylnie reprezentujących stan duszy kompozytora, kolęda z melodią przetworzoną muzycznie, ale bez trudu rozpoznawalną, to kwintesencja skojarzeń, nastrojów i uczuć, jakie są zespolone ze świętami Bożego Narodzenia. Chopin przelał w sposób genialny swoje bożonarodzeniowe uczucia na utwór. Wiele mówi to, że utwór napisany został z okazji pierwszych świąt, które kompozytor przeżył poza krajem. „Powrót” symboliczny, ale jakże wymowny. Ale właściwie do czego?

O emocjach jako rdzeniu czy istocie każdej uprawianej i odbieranej sztuki będziemy jeszcze mówić. Kończąc bożonarodzeniowe wprowadzenie, zapytajmy retorycznie: co nas pcha do świętowania Bożego Narodzenia? Co to może być, jeśli nie jakiś mniej lub bardziej zdefiniowany, skorelowany z naszymi świątecznymi uczuciami, mózgowy układ nagrody, podobny do tego, który uruchamia się zarówno w chwili

zaspokojenia pragnienia, zaspokojenia łaknienia, ale także przy takich uczuciowych stanach jak uzyskanie satysfakcji ze wzrostu pozycji społecznej, ekonomicznej, spełnienia seksualnego itp. Czy ten układ mózgowy, który gna nas czasem setki czy tysiące kilometrów, aby choć przez ten jeden dzień w roku poczuć się szczęśliwym, szczęściem właśnie dziecka, radością spotkania, być może jest tożsamy z układem, który sztucznie mogą uruchamiać: alkohol, narkotyki, pyszne jedzenie...? Chcemy takie uczucia przeżywać jeśli nie w „realu”, to przynajmniej w wirtualnym świecie, w marzeniach, we wszelkiego rodzaju „namiastkach” tych wszystkich najgłębszych, jak i najprostszyc radości... Okazuje się, że z jednej strony nie różnimy się w tym od (innych) zwierząt, nawet tych, które uważamy za prymitywne, a z drugiej strony mamy coś z „niespokojnego serca” o którym mówi św. Augustyn.

Makro- i mikroemocje – próba rozróżnienia

Reakcje emocjonalne, do których odnoszą się powyższe punkty A i B, można określić mianem „makroemocji”, które są powszechnie znane jako przejawy uczuć i których istnienie jest w jakiś sposób intuicyjnie akceptowane. Mają też swoje nazwy (strach, żądza itp.), jakkolwiek zakres tych nazw może już być dyskusyjny.

Przejawy emocji dotyczą nie tylko tych „wielkich” emocji. W ciągu jednego dnia nie odczuwamy tylko powiedzmy: trzech chwil strachu, pięciu chwil zasmucenia, dwóch chwil radosnego uniesienia, pięciu chwil pożądania i jednej chwili satysfakcji/spełnienia... A pozostały czas to co? „Pustka” emocjonalna? Stan „wyzerowania”? Neutralność? Obojętność? Oczywiście z punktu widzenia nauk interdyscyplinarnych (zajmujących się powiązaniem zachowań z systemem nerwowym) tak nie funkcjonujemy. Nie ma ani chwili, żeby w sposób świadomy lub nieświadomy falowała w nas cała gama-spektrum mikronapięć, mikrorozwiązań („neutralizacji”), czyli nie ma ani chwili, aby nie tworzyły się (i znikwały) mikroemocje... To chyba najwłaściwsza nazwa, choć pierwotnie głównie odnoszona do chwilowych grymasów na twarzy.

Postawiłbym tezę, że – posługując się analogią chemiczną – nie są to niepodzielne „atomy” – składniki bardziej złożonych emocji, z których określonych kombinacji mogłyby powstawać wspomniane wyżej

makroemocje i którym można przypisać nazwy. Przywołując jeszcze inną analogię, nie jest to też sytuacja jak w tzw. systemie Ekmana i Friesena, kodującym określone wyrazy emocji na twarzy, a wynikające z sumy (dosłownie) skurczów mięśni twarzy. Składnikami „systemu” kodowania emocji Ekmana i Friesena są poszczególne osobne anatomicznie, a nawet rozdzielone funkcjonalnie mięśnie twarzy (wspomniani autorzy wyróżnili w obrębie mięśni twarzy 46 „jednostek akcyjnych”, których skurcze „koduja” ekspresję określonych emocji – tzw. Facial Action Coding System). „Atomiczne” mikroemocje towarzyszą nam nieprzerwanie w każdej chwili życia (*notabene*, przeżywamy emocje – także makroemocje – również w czasie snu...). Mikroemocje wyobrażam sobie jako rodzaj fragmentarycznych, niepełnych makroemocji, wykorzystujących identyczne mózgowo uzwojenie, jednak uruchamiane w niepełny lub minimalny sposób, ale jednak nieprzerwanie obecne. Tworzą one rodzaj tła emocjonalnego, z którego mogą wyłaniać się „wielkie” emocje.

Istotą zarówno mikro-, jak i makroemocji są napięcia będące ciągłą grą, w jaką gra przedczołowa kora mózgu, która – jak wiadomo – nieprzerwanie ocenia prawdopodobieństwa warunkowe zdarzeń, jak i wykonywanych i planowanych przez nas działań i ich następstw na podobieństwo tzw. komputera Bayesa.

Te napięcia muszą powstawać np. jako efekt konfrontacji między predykcją i realnym efektem, między tym, co oczekiwane (upragnione), a tym, co wygląda na prawdopodobne.

„Pionkami” gry są niezliczone i często nieuświadomione asocjacje między nie tylko werbalnie możliwymi do opisanego zdarzeniami, ale też niewerbalnymi sygnałami czy czynnikami, jak zapach (określony czy nieokreślony), kolor, kształt, otoczenie w przestrzeni dźwiękowej czy materialno-przedmiotowej.

Nie musimy poszukiwać mikroemocji w spektaklu teatralnym czy widowisku sportowym. Tam oczywiście też są i falują w każdej sekundzie, a nawet ułamku sekundy, np. gdy obserwujemy lot piłki (trafi w bramkę czy nie trafi, zmieści się w linii wyznaczającej pole kortu tenisowego czy nie zmieści itp.). Mikroemocje dotyczą znacznie bardziej przyziemnych sytuacji (np. czy zdążę na autobus). Warto zauważyć, że emocje zarówno w ekspresji, jak i odbiorze (percepcji) są zarówno świadome, jak i nieświadome. Wspomniane mikroemocje prawdopodobnie głównie są nieświadome? O ekspresji i percepcji emocji będzie dalej.

Chyba jednym z najbardziej nasyconych emocjami, w tym mikroemocjami, obszarem jest każdy przejaw sztuki. Można bronić tezy, że celem każdego przejawu sztuki (literackiej, muzycznej, plastycznej) jest w rzeczywistości gra emocjami (na emocjach, z wykorzystaniem emocji).

Budowane przez twórcę napięcia są zarówno makroemocjonalne (np. dotyczą intrygi miłosnej lub znalezienia sprawcy, który dokonał zbrodni), jak również mikroemocjonalne (znajdują wyraz w obrębie frazy wypowiedzanego zdania czy gestu). Dla mnie najbliższe są przykłady muzyczne. W dziele muzycznym gra makroemocjonalna jest raczej mało wyraźna, ew. dostrzegalna może być u głęboko zaznajomionego z muzyką melomana. Ale niemal każda nuta, a na pewno nawet drobna fraza, zawiera (mikro)napięcie zazwyczaj związane z tonalnością muzyki (a ściślej – z chwilowym odchodzeniem od tonalności, które owocuje pragnieniem powrotu toniki). Modelowym przykładem takiego wywołanego u słuchacza mikronapięcia jest w muzyce sekwencja akordów zwana kadencją. Powstaje w niej napięcie związane z oczekiwaniem (przez słuchacza, a tym samym przez jego mózg) określonego rozwiązania (z reguły powrotu toniki). Kompozytorzy od zawsze grają tymi mikroemocjami i nie chodzi tylko o kompozytorów muzyki współczesnej, pełnej dysonansów i jakby zaprzeczenia „normalnej”, czyli kulturowo oczekiwanej sekwencji nut czy akordów. W tę „grę” mikronapięć emocjonalnych grali i grają wszyscy od początku muzyki, kiedy odróżniła się od rytmu. Napięcia związane z odchodzeniem i dochodzeniem do toniki/tonalności obecne są zarówno u Presleya, jak i Chopina. Badania nad tzw. *event related potentials* (ERP) z użyciem elektroencefalografu wspomaganego analizą komputerową wykazały, że mózgi nawet kompletnie nieprzygotowanych muzycznie dzieci reagują odmiennym wzorcem aktywności elektrycznej kory mózgu (co odzwierciedla aktywność neuronalną) w zależności, czy słuchana fraza muzyczna kończyła się w sposób „oczekiwany” czy też „zaskakujący” (Koelsch, 2013). Wnioski/tezy, które zaproponowałbym na tle powyższych rozważań, są następujące: warto zamiast dedukcyjnych prób definiowania pojęcia „emocji” zastosować indukcyjne, behawiorystyczne podejście, a sprowadzające się do interdyscyplinarnej analizy zachowań, reakcji emocjonalnych w określonych warunkach. Zamiast podejmować próby odpowiedzi na pytania w rodzaju „czym jest dana forma emocji?”, spróbować znaleźć odpowiedź na pytanie: „jakie są jej przejawy?”.

Komparatystyczna neurobiologia emocji

Wspomniałem już, że przejawy jakiejś formy emocji można dostrzec nawet u tak prymitywnych zwierząt, jak bezkręgowce, czy jeszcze prymitywniejszych, jak jednokomórkowce (kluczowa jest tu kwestia definicji). Może dobrym określeniem w takich wypadkach byłoby użycie terminu „protoemocje”, a nawet „preprotoemocje”? Kategoryzacje te można powiązać ze stopniem różnorodności form ich ekspresji, stopniem złożoności mechanizmów molekularnych, komórkowych czy układów funkcjonalnych komórek (najbardziej złożony jest układ nerwowy człowieka). Ale ograniczając się do świata zwierzęcego bliższego człowiekowi, bardzo płodną, a z pewnością przemawiającą do wyobraźni, jest koncepcja „trójjedynego” mózgu (*triune brain*) autorstwa Paula MacLeana. Zakłada ona współistnienie w mózgu człowieka trzech „pięter”, które nazwał odpowiednio mózgiem „gadzim” (z grubsza odpowiada to tzw. pniowi mózgu, mózdkowi, jądra podstawy, wzgórzom, opuszkom węchowym), „ssaczym” (odpowiada to strukturom limbicznym, czyli zlokalizowanym w centralno-przyśrodkowym obszarze półkul mózgowych), oraz mózgiem „naczelno-człowieczym” (płaszcz kory mózgu)...

Koncepcja MacLeana była i jest przedmiotem krytyki, zwłaszcza co do jej komparatystycznych założeń anatomiczno-funkcjonalnych. Jednak jeśli nawet wymienione wyżej anatomiczne struktury niekoniecznie ściśle odpowiadają poziomom złożoności, jakie osiągnęły mózgi wymienionych trzech kategorii zwierząt (szczegóły anatomicznych struktur mózgu, którym przypisujemy „rolę emocjonalną” będą opisane dalej), to niewątpliwa inspiracja, jaka płynie z tego modelu „trójjedynego” mózgu, polega choćby na uświadomieniu nam wszystkim, że model struktur mózgowia, nawet tych „prymitywnych” zwierząt, jak gady, nadal tkwi w nas – w każdym z nas! W każdym z nas tkwią zatem te same (takie same? czy przynajmniej analogiczne) uzwojenia (struktury neuronalne i ich układy połączeń), jakimi dysponowały i dysponują gady! Tym, co nas, ludzi, zasadniczo różni, jest przede wszystkim wykształcenie potężnego tzw. płaszcza kory mózgowej (używa się tu czasem słowa *pallium*). Mamy prawo sądzić, że zadaniem wspomnianego rozległego płaszcza kory mózgowej jest kontrola i supresja tych „niższych” mózgow ssaczych i gadzich, a w szczególności właśnie prymitywnych (?) emocji i ich przejawów. Przy czym „kontrola” to nie tylko „supresja”, czyli „hamowanie”,

ale także – a może nawet przede wszystkim – ogromne zniuansowanie naszych obserwowalnych społecznie reakcji emocjonalnych oraz zwiększenie ich dynamiki umożliwiającej nawet gwałtowne zmiany „protoemocji” w ciągu minut czy może nawet sekund. W rzeczywistości „supresja” gadzich emocji może u człowieka przeistoczyć się w sytuację odwrotną, czyli augmentację „zachowań emocjonalnych”. Jako skrajny i być może nieco retoryczny przykład niech posłuży konstatacja, że „rozwścieczone” zwierzę, takie jak krokodyl, ale też pies czy bawół, może nawet spowodować śmierć drugiego zwierzęcia (w tym człowieka), ale cóż powiedzieć o makroemocjonalnie czy mikroemocjonalnie „rozwścieczonym” człowieku uzbrojonym w karabin maszynowy...?

Najcenniejszy wniosek wynikający z modelu trójjedynego mózgu polega moim zdaniem na uświadomieniu, że te „prymitywne” mózgi ssacze i gadzie są jakoś obecnie w naszych mózgach (choć niedoświadczalnie jako obce mózgi), w mózgach każdego z nas! Nie można lekceważyć tej obecności. *Pallium* kory chroni i kontroluje te „gadzie-ssacze” mózgi i, co za tym idzie, także „gadzie-ssacze” emocje, ale są chwile (każdy je miewa!) w których świadomie czy nie do końca świadomie kontrola *pallium* bywa zdejmowana, a nawet zamiast supresji pojawia się augmentacja...

W czasie wykładu dla studentów przytaczam tu przykład podobieństwa anatomicznej pozycji (ułożenia) ciała w postaci tzw. lordozy, jaka u samic ssaków (np. szczurów) pozwala odstłonić okolice zewnętrznych narządów płciowych, co stanowi wysyłany samcom sygnał-informację o gotowości samicy do odbycia spółkowania. Wręcz szokująco na tym tle jest podobna dość typowa erotyczna poza lub pozy, jakie często obserwujemy choćby w różnych reklamówkach z udziałem atrakcyjnych kobiet reklamujących intymną bieliznę czy erotyczne sceny z filmów... W tym miejscu nie da się nie wspomnieć o tym, że – jak dobrze wiemy – istnieje niezwykle ważna np. w sukcesie zawodowym tzw. mowa ciała. A świadoma mowa ciała to też jest jedna z form ekspresji makroemocji (o formach ekspresji emocji oraz ich percepcji będzie mowa dalej). Zadam pytanie do własnego przemyślenia: jaka informacja jest (może być) sygnalizowana na fotografii modelki prezentującej pozę, która w postaci bezpośrednio danej co prawda nie prezentuje okolic krocza, ale której prosta geometryczna transformacja, np. przez obrót postaci o kąt 90 stopni (transformacja dokonywana mentalnie i prawdopodobnie

bardzo często, a może nawet najczęściej nieświadomie), taką możliwość sufluje? Powtórzę tylko: powszechnie panuje zgoda, że mowa ciała jest nośnikiem często kluczowych sygnałów dla zrozumienia czyjejś intencji lub choćby stanu emocjonalnego. Na marginesie, nie od rzeczy tu wspomnieć choćby tylko o tak licznych analizach przybieranych przez polityków postaw, reakcji czy gestów, które niekiedy zdradzają więcej na temat ich stanu emocjonalnego i prawdziwych intencji niż treść ich wypowiedzi, której często przeczą.

Pragmatyka emocji

Prawidłowa ocena i kontrola reakcji emocjonalnych, własnych i u innych osób, decyduje o osobniczym sukcesie społecznym. W tym kontekście wspomina się o tzw. inteligencji emocjonalnej. Nieprawidłowe reakcje emocjonalne cechują liczne schorzenia psychiatryczne oraz... są przyczyną schorzeń psychosomatycznych.

Nieoczekiwane, pozornie niezrozumiałe dla samego siebie reakcje emocjonalne są (mogą być) także podstawą „duchowości”, niekiedy interpretowane jako „podszept złego ducha” albo przeciwnie – „Ducha Świętego”...? Jest to oczywiście teza kontrowersyjna, wykraczająca poza aktualną metodologię prowadzenia badań nad emocjami i reakcjami emocjonalnymi. Ale raczej niekontrowersyjne jest stwierdzenie, że to emocje stoją zwłaszcza za decyzjami podejmowanymi jak gdyby wbrew analizie racjonalnej (cokolwiek by miało to znaczyć). W literaturze niejednokrotnie podkreśla się, że bardzo często tzw. wybory racjonalne to tylko pozory. Czy w ogóle rzeczywistość jest antynomia osoba racjonalna vs osoba nieracjonalna i stojąca u jej podstaw antynomia: mózg racjonalny vs emocjonalny? Jakikolwiek tzw. wybór racjonalny obciąża świadomość ograniczeń dostępnej wiedzy (prawdziwość przesłanek) oraz niepewność co do właściwej ich interpretacji, jak i logicznej poprawności rozumowania. Tymczasem sam dobór „racjonalnych” przesłanek jest (w dużej mierze ukrycie) obciążony emocjonalnie (i nieświadomy). Jakiegokolwiek wahanie (w racjonalnym wyborze) jest przezwyciężane przesłanką emocjonalną (np. tak mi będzie przyjemniej...).

Pragmatyczno-społeczna rola emocji jest kluczowa dla tworzenia lub utraty więzi międzyludzkich, dla zdolności/niezdolności do współpracy

w obrębie „celowej grupy”, np. drużyny sportowej czy załogi samolotu itp., dla oddziaływania reklam, ma także wpływ na wybory polityczne itp.

Podobnie za fałsz należy uznać rzeczywisty czysto racjonalny wybór, fałszem jest też „unifikująca duchowość”, zwłaszcza egzemplifikowana w wypowiedziach takich jak „duch narodu”, „duch drużyny” itp. Tym, co nazywamy „duchem narodu” czy może nawet wręcz „narodem”, jest grupa ludzi, która podobnie emocjonalnie reaguje na treści i symbole takie jak hymn państwowy czy inna patriotyczna pieśń, godło, barwy narodowe, bohaterskie postaci i czyny itp. Osobiście nazwałbym Polakiem każdego, kto wykazuje wegetatywne „pozytywne” reakcje (przyspieszone bicie serca, łzy w oczach itp.) na ww. wspomniane treści czy symbole. Słowo „wegetatywne” ma tu szczególne znaczenie, ponieważ to właśnie reakcje wegetatywnego układu nerwowego należą do wolicjonalnie albo niekontrolowalnych, albo bardzo słabo podlegających kontroli, co oznacza, że w powszechnym rozumieniu to one wyrażają „prawdę” o naszych emocjonalnych więziach. To jest to, co od wieków nazywa się wspomnianym wcześniej „sercem” (i przeciwstawia „rozumowi”).

Powyższe zagadnienia wiążą się z ekspresją i komunikacją naszych emocji, ale także z ich umiejętnością odczytania (percepcji), wczuwania się w drugą osobę i jej stan emocjonalny.

Emocje komunikujemy oraz odczytujemy poprzez nasze reakcje emocjonalne. Zarówno komunikacja (równoważnikiem jest „ekspresja”), jak i odczytywanie (czyli „percepcja”) bywają świadome (uświadomione) i nieświadome, a ponadto mogą być wolicjonalnie kontrolowane lub niekontrolowane (lub tylko częściowo kontrolowane).

Ekspresja emocji dokonuje się z wykorzystaniem następujących zasadniczych środków:

- a. reakcje trzewnego ruchowego układu nerwowego (autonomicznego) – od strony anatomicznej, czyli nieobserwowalnej gołym okiem,
- b. mowa (prozodia, jak i sama treść!),
- c. mimika twarzy,
- d. mowa ciała (*body language*) i bardziej ogólnie komunikacja niewerbalna (w tym m.in. sposób ubierania się, dotyk).

Percepcja emocji również odbywa się w sposób uświadomiony i nieświadomy. Nieświadomy aspekt percepcji emocji jest związany (w dużej mierze) z tzw. neuronami zwierciadlanymi (Rizzolatti, 2004; Gładziejewski, 2011). Są to neurony, które kodują uruchomienie podobnych reakcji

motorycznych i wegetatywnych u, nazwijmy to, „obserwatora”, jakie występują u osoby obserwowanej („obserwabli” – propozycja terminu rodem z fizyki kwantowej). Przykładowo obserwując kopnięcie w okolicę krocza, doznajemy (możemy doznać) niejako odruchowego przyjęcia postawy chroniącej tę okolicę. Neurony zwierciadlane są najprawdopodobniej materialnym podłożem (?) empatii, czyli mniej lub bardziej silnie doznawanego współuczucia tych samych lub podobnych doznań jak u osoby „obserwowanej”. Empatia (możliwość empatii) cechuje nie tylko człowieka, ale także przynajmniej niektóre naczelne, a niewykluczone, że również wiele innych zwierząt o „niższym” stopniu rozwoju.

Refleksja systematyczna z wnioskami i rekomendacjami

Od czasów Paula Broca i Jamesa Papeza poznano wiele podobnych „ośrodków”, ale nadal w rozległych obszarach mózgu, w tym w szczególności kory mózgowej, nie potrafimy wskazać czegokolwiek konkretnego i pewnego odnośnie do ich funkcji. Obwód Papeza jest już jedynie traktowany jako element historii w poznaniu mózgu, ale częściowo pokrywa się także ze strukturami zwanymi układem limbicznym, co do których panuje zgoda, że pełnią ważną funkcję w kontroli emocji. *Nota bene* „kontrola” to ostrożny termin, unikający deklaracji co do mózgowego źródła „powstawania” emocji – czy stanów emocjonalnych. Jednak nie uważamy, że cały układ limbiczny odgrywa rolę w emocjach. Do układu limbicznego zalicza się bowiem również zakręt parahipokampalny z hipokampem oraz ciałka suteczkoweate podwzgórza i sklepienie, ale ich związek z emocjami jest mniej jednoznaczny.

Natomiast uważa się, że szczególnie istotną rolę w kontroli emocji grają:

- a. zakręty nadoczodołowe oraz przyśrodkowa kora przedczołowa,
- b. brzuszna część jąder podstawy (jądro półleżące – *nucleus accumbens*),
- c. przyśrodkowo-grzbietowe jądro wzgórza – *nucleus mediodorsalis thalami* (czyli inne niż jądra przednie wzgórza wymienione w obwodzie Papeza),
- d. ciało migdałowe,

- e. zakręt obręczy,
- f. jądra przegrody (*septal area*),
- g. część podwzgórza (*hypothalamus*).

Należy zwrócić uwagę, że ww. struktury należałyby właściwie wymienić w liczbie mnogiej, gdyż wszystkie występują w obu półkach mózgu, ale utrzymamy „manierę” stosowania liczby pojedynczej, chyba że użycie liczby mnogiej będzie istotne. Ponadto zwraca się również uwagę na rolę w mechanizmach emocji takich obszarów jak substancja szara okołowodociągowa (*periaqueductal gray* – PAG) czy pole brzuszne nakrywki (VTA – *ventral tegmental area*) – pierwszy z nich kontrolujący przewodzenie bodźców bólowych, drugi grający ważną rolę w tzw. obwodzie nagrody (zob. dalej).

Spośród ww. najwięcej uwagi poświęca się ciału migdałowatemu będącemu relatywnie dużym skupiskiem kilku jąder o łącznej średnicy nieco ponad 1 cm. Nie ma wątpliwości, że jest ono swego rodzaju centrum („huba”) w tworzeniu asocjacji zdarzeń natury mentalnej i czysto percepcyjnej, w czym upatrujemy tworzenie i gromadzenie „pamięci emocjonalnej”, oraz w nadawaniu zdarzeniom wartościowości (*valence*) i wydatności (*salience*). Centralna rola ciała migdałowatego wynika z anatomii jego połączeń. Otrzymuje ono z jednej strony kolaterale od wszystkich dróg i układów czucia (w tym warto podkreślić anatomiczne struktury czucia węchu, znajdujące się w bezpośrednim sąsiedztwie ciała migdałowatego. Intuicyjnie możemy założyć, że węch posiada w emocjach ogromne znaczenie (zwłaszcza u zwierząt). Z drugiej strony ciało migdałowate ma rozległe połączenia z całym układem limbicznym, a połączenie z podwzgórzem, niewątpliwie kluczowe w komunikacji i zarządzaniu układem hormonalnym oraz aktywnością układu wegetatywnego, ma charakterystyczną postać zwaną prążkiem krańcowym (*stria terminalis*).

To wielokierunkowe połączenie ciała migdałowatego posłużyło współczesnemu neuronukowcowi Josephowi LeDoux do wysunięcia koncepcji „drogi wysokiej” i „drogi niskiej” aktywacji ciała migdałowatego (LeDoux, 2001). Droga „niska” to droga „bezpośredniej” aktywacji przez doznania czuciowe wszelkich modalności, a droga „wysoka” oznacza przejście doznania przez korę. Droga niska umożliwia niemal błyskawiczne uruchomienie odpowiedzi, które czasem może skutkować błędem wywołanym zbyt szybką reakcją, ale w niektórych może

uchronić przed najgroźniejszymi konsekwencjami, w tym śmiertelnymi. Jako przykład podawana jest odruchowa reakcja odskoku w odpowiedzi na nagłe poruszenie się węża do spryskiwania trawnika, mylnie potraktowanego jako nagły ruch niebezpiecznego gada... „Droga wysoka” przez pośrednictwo kory mózgowej uchroni nas przed „ośmieszeniem”, ale gdyby jednak „wąż” okazał się prawdziwym niebezpiecznym gadem, nasze zdrowie, a nawet życie mogłoby być wystawione na ogromne niebezpieczeństwo. Ale zwróćmy uwagę, że nie istnieje możliwość uruchomienia tylko jednej z tych (konceptyjnych) dróg. Jeśli nawet okażemy stoicki spokój (niejako będący rezultatem kognitywnej analizy bodźca przez korę mózgu), to i tak najprawdopodobniej już wcześniej ciało migdałowate, pobudzone „niską drogą”, wyśle do podwzgórza sygnał, aby pobudzić układ wegetatywny oraz korę nadnerczy i podnieść poziom adrenaliny...

Warto teraz przypomnieć sobie, o czym była już mowa, a mianowicie, że jeśli zwierzę ma dostęp do powtórzenia osobistego doświadczenia (zjedzenie czegoś, kopulacja itp.), konieczne jest zapamiętanie go choćby poprzez „prostą” asocjację wartościowości (pozytywna-negatywna, przyjemna-nieprzyjemna itp.) z epizodem. Warto dodać, że u zwierząt także obserwujemy pamięć epizodyczną, mimo że nie potrafią przekazać jej werbalnie. Wiadomo, że co najmniej jednym z neurobiologicznych mechanizmów wykorzystywanych przez neurony ciała migdałowatego (zwanego także amygdalą) jest tzw. długotrwałe wzmocnienie synaptyczne (*long term potentiation* – LTP) m.in. z użyciem sygnalizacji glutamatergicznej i receptora glutaminianu typu NMDA. Pośrednim dowodem na udział receptora NMDA jest to, że podawanie jego antagonistów blokuje reakcje warunkowania strachu. Amygdala jest innymi słowy asocjacyjnym spoiwem, które tworzy kolektywnie coś, co nazywamy „pamięcią emocjonalną”. Może być ona świadoma, ale bardziej intrygujące wydają się jej aspekty nieuświadomione, np. jakiś bardzo nieprzyjemny epizod typu pogryzienie przez psa w dzieciństwie, które może być „zapomniane” w sensie utraty pamięci epizodycznej, tym niemniej przez całe życie może być pod względem emocjonalnym trwałe, co będzie się odzwierciedlać w postaci większego niż przeciętnie lęku przez psami...

Obwód Papeza odegrał jako neurologiczna koncepcja istotną rolę, ponieważ zapoczątkował zasadniczo aktualną również w chwili obecnej

metodę badań czy podejście do wyróżniania wielu innych „obwodów” funkcjonalnych połączeń w mózgu jako materialnych „substratów”, jeśli nie „tworzenia”, to przynajmniej „kontroli” mózgowopochodnych zjawisk emocjonalnych. Współcześnie opisuje się więc np. obwód strachu (*fear circuitry*) czy obwód nagrody (*reward circuitry*). Zalicza się do nich także elementy niewymienione wyżej (ppkt. a-g). I tak do „obwodu nagrody” (*reward circuitry*) – będącego zasadniczym „wzmacniaczem” (*reinforcer*) tego, czemu ciało migdałowate „przypisze” pozytywne (przyjemne) znaczenie, zalicza się: oprócz ciała migdałowatego, kory przedczołowej i jądra półleżącego także VTA.

Dwie ostatnie z wymienionych struktur czyli jądro półleżące i VTA, a zwłaszcza VTA, są warte poszerzonej wzmianki i dalszych badań interdyscyplinarnych. Jądro półleżące to dość duże zgrupowanie neuronów (ogólnie „jądrami” w mózgu nazywamy wszelkie wyróżnione anatomicznie ugrupowania neuronów poza korą mózgu i mózdzku), leżące w dolnej przedniej części półkul mózgowych (tzw. brzuszne prążkowie). VTA z kolei jest dość niewielkim, z trudnością i tylko w mikroskopie rozpoznawalnym jądrem w śródmózgowiu. Leży ono blisko istoty czarnej i podobnie jak część „zbita” istoty czarnej (*pars compacta*) jest utworzone z neuronów, które jako neurotransmitera, czyli substancji wydzielanej przez synapsy w celu komunikacji z innym neuronem, używają dopaminy. Aksony neuronów VTA prowadzą do jądra półleżącego i tam przy ich zakończeniach syntetyzują i wydzielają dopaminę. Dość powszechna jest opinia, że dopamina jest „hormonem szczęścia”. To tylko częściowa prawda, bo dopamina odgrywa inne role w innych strukturach mózgu (np. kontroluje wydzielanie prolaktyny) ale rzeczywiście stymulacja wydzielania dopaminy przez neurony VTA w synapsach, jakie tworzą z neuronami jądra półleżącego, jest – jak można się domyślić – istotnym funkcjonalnym elementem „obwodu nagrody” (bo połączenie VTA i *nucleus accumbens* jest częścią anatomiczną tego obwodu). Podsumowując, obwód nagrody uważany jest za podstawową strukturę, której aktywacja oznacza stan przyjemności, ale może lepiej określić go stanem „chcianym”, apetytywnym.

W pewnym sensie przeciwstawnym do „obwodu nagrody” jest obwód strachu (*fear circuitry*), który można utożsamiać z bodźcami „awersyjnymi”. Zalicza się do niego wspomniane wyżej struktury limbiczne, ale także hipokamp oraz jądro miejsca sinawego (*locus*

coeruleus) – charakterystyczne „ciemno” podbarwione i łatwo identyfikowalne, choć bardzo małe jądro w pniu mózgu blisko dna komory IV z komórkami nerwowymi, których neurotransmiterem jest noradrenalina.

Ostatecznie próbując odpowiedzieć na pytanie, „gdzie (w mózgu) są emocje?”, można odnieść wrażenie, że znamy dość dokładnie struktury i ich połączenia, które „obsługują” mózg pod względem „emocjonalności”, ale to efekty działania całego mózgu dają wyraz emocjom. Co więcej, na zasadzie czegoś, co przypomina sprzężenie zwrotne działania mózgu (i z pewnością efekty tych działań), oddziałują na nasz stan emocjonalny właściwie w czasie rzeczywistym. Gra aktorska „udająca” określone emocje nie pozostaje obojętna dla grającego aktora. Nie da się odizolować emocji od innych funkcji mózgu. Nie da się (przywołując koncepcję LeDoux) działać tylko drogą „wysoką” czy tylko drogą „niską”. A paradoks tkwi także w tym, że nawet ostentacyjne blokowanie wyrazów emocjonalności będzie interpretowane jako niosące informacyjne emocjonalne znaczenie...

Na zakończenie „bożonarodzeniowego” wykładu o interdyscyplinarnie rozumianej neurobiologii emocji powracam do wyjściowego motywu świątecznego. Natura tej szczególnej emocjonalnej mocy tkwi we wspomnianej uprzednio empatii, ale już nie międzyludzkiej empatii, ale empatii między człowiekiem i Bogiem-dzieckiem. Niesie to paradoks. Bóg niby wszechmocny, lecz jako dziecko i jak dziecko potrzebujący opieki, troski... To temat na zupełnie inne rozważania, natomiast jeśli „wszechmoc” emocji sprowadzić do ich niekwestionowanej roli w naszym życiu, to właśnie emocje okazują się w wielu przypadkach jeśli nawet nie „wszechmocne” w indywidualnym i zbiorowym życiu człowieka i społeczeństwa/narodu, to z pewnością – jeśli chodzi o mózg – są wszechobecne, o czym świadczą badania nad neuronalnym podłożem reakcji emocjonalnych. Ich neuronalne korelaty, patrząc od strony funkcjonalnej, są właściwie wszechobecne w mózgu, a w jego działaniach są niczym barwniki, które nieustannie podbarwiają „strumień świadomości” w szerokiej gamie różnorodności i intensywności od czarującej migotliwej tęczy, aż do całkowitej utraty przejrzystości i zamęcenia.

BIBLIOGRAFIA

- Augustyn, św. (1987). *Wyznania*. Przeł. Z. Kubiak. Warszawa: PAX.
- Bear, M.F., Connors, B.W., & Paradiso, M.A. (2016). *Neuroscience. Exploring the Brain*. Philadelphia: Wolters Kluwer.
- Bremer, J. (2005). *Jak to jest być świadomym. Analityczne teorie umysłu a problem neuronalnych podstaw świadomości*. Warszawa: Wydawnictwo IFiS PAN.
- Cannon, W.B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, 39, 106–124.
- Ciccarelli, S., & White, J.N. (2023). Motywacja i emocje. W: W. Domachowski (Red.), *Psychologia*. Przeł. A. Bukowski, & J. Środa. Poznań: Dom Wydawniczy Rebis, 346–379.
- Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience*, 5, 583–589.
- Davis, K.L., & Panksepp, J. (2017). *The Emotional Foundations of Personality. A Neurobiological and Evolutionary Approach*. London: W.W. Norton & Company New York, London.
- Gładziejewski, P. (2011). Czy empatia jest symulacją mentalną? Dyskusje z podejściem reprezentacjonistycznym w koncepcji neuronów lustrzanych. *Diametros*, 27, 108–120.
- Koelsch, S. (2013). *Brain and music*. Chichester: Wiley-Blackwell.
- LaPointe, L.L. (2012). *Paul Broca and the Origins of Language in the Brain*. San Diego: Plural Publishing.
- LeDoux, J. (2001). *Mózg emocjonalny. Tajemnicze podstawy życia emocjonalnego*. Warszawa: Media Rodzina.
- Papez, J.W. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 38(4), 725–743. DOI: 10.1001/archneurpsyc.1937.02260220069003.
- Purves, D., Augustine, G.J., Fitzpatrick, D., Hall, W.C., LaMantia, A.-S., McNamara, J.O., & White, L.E. (Red.) (2008). *Neuroscience*. Sunderland: Sinauer Associates Inc.
- Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27, 169–192.

Mateusz Jarmużewski
Windesheim University of Applied Sciences
<https://orcid.org/0000-0002-8398-1398>
<https://doi.org/10.35765/slowniki.482>

Neurobiologiczne podstawy odpowiedzialności moralnej

Streszczenie

DEFINICJA POJĘCIA: Odpowiedzialność za działanie jest podstawową prerogatywą człowieka jako istoty moralnej. W większości systemów etycznych warunkiem odpowiedzialności moralnej jest (dobro)wolność działania, ściśle związana z pojęciem wolnej woli. Aktom moralnym dokonywanym w świecie fizycznym – jak i samej odpowiedzialności – można przypisać naturalne podstawy i prawidłowości. Zgodnie z obecnym stanem wiedzy są to przede wszystkim wyjaśnienia na poziomie neurobiologicznym.

ANALIZA HISTORYCZNA POJĘCIA: Biologiczne podstawy zachowań były przedmiotem zainteresowania od starożytności. Świadomość ta pozostawała widoczna w zachodniej antropologii i etyce chrześcijańskiej, jednak neurobiologiczne uzasadnianie (braku) odpowiedzialności moralnej należy już do epoki nowożytnej. Rozpowszechniony w tym kontekście sceptycyzm, m.in. wobec idei wolnej woli, mimo że nie zawsze pociąga za sobą zanegowanie odpowiedzialności, wyraźnie stawia pod znakiem zapytania jej naturalne podstawy.

UJĘCIE PROBLEMOWE POJĘCIA: Niektóre nowe trendy neuronauk i biologii systemowej mogą stanowić wyjście dla filozoficznego impasu w dyskusji nad wolną wolą, w kontekście tak deterministycznych, jak i stochastycznych wyjaśnień neuronauk poznawczych. Trendy te naświetlają niebezpośredni, współzależny i długoterminowy charakter autonomii organizmu jako całości, co ma kluczowe znaczenie dla rozumienia zarówno podmiotu, jak i przedmiotu odpowiedzialności moralnej.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI:

Analiza neurobiologicznych podstaw ludzkiego działania sugeruje rozproszone, a jednocześnie szerokie spektrum odpowiedzialności moralnej. Tak rozumiana odpowiedzialność obejmuje całość neurobiologicznego dynamizmu osoby, co pozostaje zakorzenione w tradycyjnej antropologii personalistycznej, ze szczególnym uwzględnieniem katolickiej tradycji teologiczno-moralnej oraz społeczno-etycznego kontekstu współczesności.

Słowa kluczowe: neurobiologia, odpowiedzialność moralna, wolna wola, etyka cnót, działanie moralne

Definicja pojęcia

Kwestia odpowiedzialności, szczególnie w kontekście normatywno-etycznym, pozostaje kluczowa nie tylko jako uzasadnienie moralnego wartościowania na poziomie osobowym, ale także jako podstawowa prerogatywa człowieka jako istoty moralnej. Odpowiedzialność za szeroko pojęte działanie (ang. *responsibility*) – lub, mówiąc ściślej, możliwość przypisywania tak pojmowanego działania (wraz z jego konsekwencjami) do konkretnej osoby ludzkiej (ang. *accountability*) – można traktować jako podstawowe założenie większości systemów etycznych. Odpowiedzialność za działanie w kontekście jego normatywnej oceny w kategoriach „dobra” i „zła” nazywamy odpowiedzialnością moralną.

Zakres działania, za które możemy być odpowiedzialni (*responsible*) – lub za które możemy być w uzasadniony sposób pociągnięci do odpowiedzialności (*accountable*) – pozostaje szeroki. Podmiotem odpowiedzialności pozostaje człowiek jako osoba, jednakże nie brakuje też etycznej refleksji nad moralną odpowiedzialnością zbiorową lub historyczną. Przedmiotem odpowiedzialności moralnej mogą być nie tylko konkretne czyny lub zachowania, lecz także postawy, intencje lub cnoty moralne. Katolicka teologia moralna oraz etyka personalistyczna traktuje w tym kontekście o tzw. aktach moralnych. Zakładają one ludzkie działanie w sensie ścisłym (*actus humanus*), w odróżnieniu od aktywności (*actus hominis*), która jedynie zachodzi w człowieku/przez człowieka (Tomasz z Akwinu, 2006, S. *Th.* I–II, 1.1).

Pomimo że do aktów moralnych zalicza się również świadome i dobrowolne myśli lub postanowienia, nie wyczerpuje to całego spektrum ludzkich aktywności i postaw podlegających moralnemu wartościowaniu, moralnej ocenie, a co za tym idzie – również odpowiedzialności moralnej. Rozumiane w szerokim sensie działanie moralne będzie więc obejmować wszystkie aspekty ludzkiej egzystencji, za które możemy (i być może powinniśmy) być odpowiedzialni.

Większość systemów etycznych, jak również spora część ludowych przekonań w ramach kultury zachodniej, podaje w tym kontekście jeden podstawowy warunek. Warunkiem tym jest (dobro)w o l n o ś ć podejmowanych działań, postanowień i zaniechań, wyrażanych postaw czy sposobów bycia. Tylko to, co człowiek czyni lub do czego dąży w sposób wolny, może być przedmiotem odpowiedzialności moralnej (Prümmer,

1956, s. 8.). Mimo że wolność jako podstawa odpowiedzialności moralnej wydaje się założeniem dość uniwersalnym, odpowiedź na pytanie, czym taka wolność właściwie jest, zależeć będzie od filozoficznych, teologicznych i kulturowych trendów i epok. To samo dotyczy pytania o stopień wolności konieczny (i wystarczający) do uzasadnienia zarówno moralnych prerogatyw działania, jak i odpowiedzialności moralnej osoby.

Z warunkiem (dobro)wolności wiąże się pytanie o przyczyny, źródła, naturę lub – mówiąc ogólnie – podstawy ludzkiego działania. Akceptując antropologiczne założenia ciągłości lub przynajmniej wyraźnego związku człowieka ze światem materialnym, każdemu ludzkiemu działaniu, myśli czy postawie można, co do zasady, przypisać pewne fizyczne, chemiczne, biochemiczne czy wreszcie typowo biologiczne przyczyny oraz leżące u ich podstaw prawidłowości. Wśród różnorodności naturalnych podstaw można wyróżnić to, co dane działanie czy postawę bezpośrednio powoduje (lub je fizycznie realizuje), oraz to, co je (bio)fizycznie umożliwia. Używając języka zaczerpniętego z teologii stworzenia, możemy tu mówić o przyczynach pierwszego i drugiego rzędu. Oczywiście tego typu poziomów można wyróżnić znacznie więcej, a relacje między nimi nie są ani oczywiste, ani łatwe do wyjaśnienia. Zadanie to staje się jeszcze trudniejsze, gdy do refleksji nad podstawami ludzkiego działania włączone zostają bezpośrednio i (przede wszystkim) pośrednio przyczyny o charakterze nadprzyrodzonym.

Jeśli, idąc dalej tropem tej analogii, powiążemy (lub w pewnym zakresie u t o ż s a m i m y) naturalne podstawy ludzkich działań z przyczynami nadprzyrodzonymi – lub ze stworzonymi prawami natury – wkroczymy w ten sposób w pole klasycznego teologicznego sporu o wolną wolę człowieka w kontekście wszechwiedzy i opatrności Boga. Nie umniejszając odrębnego charakteru wyjaśnieniom teologicznym oraz nie ulegając bezkrytycznie koncepcjom pan(en)teistycznym, należy stwierdzić, że uniknięcie płytkiego interwencjonizmu pozwala usytuować teologiczną debatę o boskiej i ludzkiej wolności w kontekście pytania o wszelkie istotne podstawy ludzkiego działania. Co szczególnie ważne, poszukiwanie bezpośrednich przyczyn i pośrednich prawidłowości rządzących ludzkim działaniem moralnym (lub – szerzej – tym aspektem ludzkiej aktywności, które mogą stać się przedmiotem oceny moralnej) dotyczy zarówno tego, co takie działanie umożliwia i powoduje, jak i tego, co uzasadnia przypisanie działającemu podmiotowi odpowiedzialności.

Pytając, czy dane działanie należy uznać za wolne, pytamy między innymi o to, (a) kto lub co stoi za podejmowanym czynem czy wyrażaną postawą, oraz (b) czy dana osoba mogła lub była w stanie postąpić inaczej. Nierzadko pytamy również o to, (c) na ile to, co człowiek robi lub co sobą reprezentuje, jest wyrazem zdolności reagowania na argumenty. Innymi słowy: czy osoba postępuje zgodnie z jej (mniej lub bardziej uzasadnionymi) racjami i czy jest je w stanie w sposób nieskrępowany rozpoznać, zrozumieć i wprowadzić w życie. Mówiąc kolokwialnie: czy osoba naprawdę „robi to, co chce” lub – w tym przypadku – to, co (akurat) uważa za słuszne. Pytania (a) i (b) wyrażają dwa podstawowe filozoficzne warunki istnienia wolnej woli, będące pod znaczną presją ze strony nauk ścisłych i przyrodniczych. Pytanie o „zdolność reagowania na racje” (c) wydaje się pod tym względem mniej skomplikowane, jednak i ono pociąga za sobą m.in. kwestię naturalnych podstaw racjonalności.

Fakt, iż moralnie działający człowiek funkcjonuje w fizycznym świecie przede wszystkim jako żywy organizm, uzasadnia skupienie uwagi na wyjaśnieniach biologicznych. Zlokalizowanie ośrodków ludzkiego myślenia i działania w centralnym układzie nerwowym, a szczególnie w samym mózgu, tłumaczy centralne miejsce szeroko pojętych neuronauk. W polskojęzycznym kontekście terminy „neurobiologia” i „neurobiologiczny” bywają używane do określenia całego obszaru badań współczesnych neuronauk. Pierwotnie jednak neurobiologia pozostaje metodologicznie dziedziną biologiczną i wykracza poza (wieloaspektowe) badanie funkcjonowania mózgu. W przeciwieństwie do podejścia czysto neuronaukowego (włączając w nie m.in. neuronauki poznawcze czy neuropsychologię) podejście neurobiologiczne uwypukla funkcjonowanie organizmu jako całości oraz biologicznie uwarunkowaną rolę, jaką w tym procesie spełnia mózg i cały system nerwowy.

Neurobiologiczne podstawy odpowiedzialności moralnej obejmują zatem procesy *stricte* neuronalne, jak i szersze, ogólnobiologiczne tło wolitywnych działań człowieka. W badaniu tej drugiej grupy procesów nieocenione są takie pokrewne dyscypliny jak biologia ewolucyjna, biologia systemowa, etologia i ekologia. Podobnie jak w wielu gałęziach neuronauk poznawczych, próba zrozumienia funkcjonowania układu nerwowego i jego interakcji z całym organizmem (w przypadku człowieka z podmiotem działania moralnego) nie może obejść się bez naświetlenia

procesów fizycznych, biofizycznych i biochemicznych. Warto wspomnieć, że poza podstawową wiedzą o biofizyce i chemii mózgu próby wyjaśnienia wolitywnych aspektów funkcjonowania organizmu na poziomie np. fizyki kwantowej zwykle pozostają w sferze hipotez i spekulacji.

Rozważania filozoficzne dotyczące determinizmu, wolnej woli oraz epistemologicznego statusu neuronauk mają znaczący wpływ na interpretację neurobiologicznych danych dotyczących genezy ludzkiego działania. Kwestia wolnej woli jako takiej – podejmowana z perspektywy metafizycznej, analitycznej, teologicznej czy z perspektywy tzw. neurofilozofii (filozofii neuronauk) – jawi się jako kluczowa tak dla właściwego usytuowania pojęcia odpowiedzialności moralnej, jak i dla możliwości nakreślenia jej naturalnych uzasadnień. Mimo to ewentualna teza o nieistnieniu wolnej woli nie zawsze skutkuje odrzuceniem odpowiedzialności moralnej (Fisher, 2007).

To samo dotyczy ściśle filozoficznego problemu umysł-mózg/ umysł-ciało czy też problemu tak zwanej sprawczości umysłowej (*mental causation*). Również rozważania na temat istnienia i natury samoświadomości oraz typowo psychologiczne opisy sprawczości i wolności nie zawsze pokrywają się z uzasadnieniami neurobiologicznymi. Odpowiedzialność moralna pozostaje też wyraźnie odróżniona od pojęcia odpowiedzialności prawnej, cywilnej czy służbowej.

Analiza neurobiologicznych podstaw tak rozumianej odpowiedzialności skupia się na wyjaśnieniach w ramach nauk przyrodniczych, traktując zdawkowo lub kontekstualnie wyjaśnienia wykraczające poza ten obszar (jak choćby wspomniane wyjaśnienia psychologiczne, fenomenologiczne czy teologiczne). Jednocześnie, ponieważ pojęcie odpowiedzialności moralnej pozostaje przede wszystkim pojęciem etycznym i społecznym – i samo w sobie wykracza poza metodologiczne ramy nauk przyrodniczych – wyniki podejmowanej analizy mieszczą się z definicji w ramach interdyscyplinarnego dialogu oraz filozofii i teologii nauki.

Analiza historyczna pojęcia

Pomimo że początków systematycznej wiedzy o systemie nerwowym (zarówno u ludzi, jak i u zwierząt) możemy dopatrywać się dopiero w XVII–XIX wieku, biologiczne podstawy ludzkich zachowań były

przedmiotem zainteresowania od starożytności. Podobnie ma się sprawa z neuronaukami poznawczymi i z naukami o poznaniu w klasycznym, filozoficznym ich ujęciu. Powszechnie uznane łącznie myślenia, emocji i decyzyjności z mózgiem jest także zdobyczą nowożytności, jednak relacja tych procesów z funkcjonowaniem ludzkiego ciała fascynowała już wczesnych etyków, metafizyków i epistemologów.

Historycznie obecna świadomość wpływu biologii na to, co – ale przede wszystkim jak – wybieramy, nie może umknąć uwadze badaczy. Jednocześnie próby neurobiologicznego uzasadnienia samej odpowiedzialności (lub jej braku) należy wiązać z nowożytnym rozwojem nauk przyrodniczych i empirycznych oraz z ich coraz bardziej radykalnym oddzieleniem od dyscyplin humanistycznych i teologicznych. Generalnie rzecz biorąc, systematyczna refleksja nad ludzką wolnością i odpowiedzialnością podążała zgoła innymi torami niż naukowe próby wyjaśnienia ludzkiego działania oraz rządzących nim prawidłowości. Taki stan rzeczy utrzymywał się mniej więcej do epoki oświecenia, a stopniowa zmiana podejścia nieprzypadkowo zbiegła się z rozkwitem m.in. współczesnej neuropsychologii, nauk o mózgu czy tradycji analitycznej w filozofii.

Jak wiadomo, starożytność (nie tylko ta zachodnia) umiejscawiała procesy poznawcze w przeważającej większości w ludzkim sercu. Prawdopodobnie ze względu na oczywisty związek z krwią, serce było powszechnie uważane za siedlisko rozumnej duszy, podczas gdy emocje i pragnienia zmysłowe bywały (jak u Arystotelesa) lokalizowane gdzie indziej (w tym przypadku w wątrobie). Jednym z nielicznych wyjątków był tu grecki Hipokrates (460–377 p.n.e.), którego poglądy na rolę mózgu nie odbiegały znacząco od tych prezentowanych współcześnie. W Aleksandrii III w. p.n.e. Herophilos po raz pierwszy rozróżnił mózdzek i dwie półkule mózgu oraz zasugerował, że to właśnie mózdzek (ze względu na większą gęstość) musi kontrolować mięśnie. Erasistratos bezbłędnie zlokalizował logiczne myślenie w korze mózgowej, spekulując, że ludzka inteligencja musi mieć związek z liczbą widocznych bruzd (Gross, 1987). Późniejsze prace Galena (130–200 n.e.) otwarcie sugerowały, że mózg (wraz z rdzeniem kręgowym jako jego przedłużeniem) jest miejscem odczuwania, myślenia oraz kontroli zachowań.

Mniej więcej taki stan wiedzy, wliczając np. opisy powstających w mózgu zaburzeń neuropsychiatrycznych autorstwa Awicenny, utrzymywał się do epoki odrodzenia. Kartezjusz (1596–1650), uważany za

pioniera badań nad mózgiem i umysłem, znany jest głównie ze swojej teorii dualizmu, sugerującej między innymi, że to szyszynka jest miejscem, w którym umysł wchodzi w interakcję z ciałem (po zarejestrowaniu mechanizmów kierujących krążeniem płynu mózgowo-rdzeniowego). Wkład Kartezjusza w neurobiologię działania jest jednak szerszy i obejmuje m.in. teorię refleksów. Tłumaczyła ona zachowanie odruchowe jako reakcję na bodźce zewnętrzne, poruszające skórę i włókna, uwalniające zwierzęcą duszę i unerwiające mięśnie (Elias & Saucier, 2006, s. 6). To mechaniczne wyjaśnienie otworzyło drzwi do badania neuronalnych podstaw zwierzęcych i – na ich podstawie – ludzkich zachowań jako takich.

Kolejny ważny etap w rozwoju neurobiologii wyznaczają odkrycia poszczególnych funkcji systemu nerwowego oraz coraz dokładniejsze wiązanie ich z anatomią. Za początek tej epoki uważa się odkrycie Bella (zm. 1842) i Magendiego (zm. 1840), ukazujące odmienne funkcje nerwów rdzenia kręgowego. Nie sposób wymienić wszystkich punktów zwrotnych coraz szybciej rozwijającej się dziedziny. Jednym z najważniejszych, z punktu widzenia współczesnych nauk o mózgu, było niewątpliwie mikroskopowe opisanie komórek nerwowych (Golgi, zm. 1926) oraz sieci ich połączeń (Ramón y Cajal, zm. 1934), a później także anatomii i funkcjonowania samych połączeń synaptycznych (Charles Scott Sherrington, zm. 1952), roli neurotransmiterów (Loewi, zm. 1969) czy wreszcie istnienia potencjałów czynnościowych (Bernstein, zm. 1917).

Druga połowa XX wieku przyniosła kolejny skok wraz z wynalezieniem skutecznych technik neuroobrazowania, takich jak rezonans magnetyczny (MRI), magnetoencefalografia (MEG) i pozytonowa tomografia emisyjna (PET), a także efektywne wykorzystanie starszej techniki elektroencefalografii (EEG) czy przezczaszkowej stymulacji magnetycznej (TMS). Skany funkcjonalne (fMRI) umożliwiły zbieranie informacji o bieżącej aktywności w różnych częściach mózgu (ze słabą rozdzielczością czasową). Z kolei zapisy EEG informują naukowców o potencjałach elektrycznych mierzonych w czasie rzeczywistym (ze słabą rozdzielczością przestrzenną).

Te obiecujące możliwości zainspirowały naukowców takich jak Benjamin Libet (zm. 2007) do badania neuronalnych podstaw ludzkiego działania oraz m.in. powszechnego doświadczenia wolnej woli. Od

połowy lat 80. przeprowadzono setki eksperymentów (głównie z wykorzystaniem EEG i fMRI) próbujących uchwycić procesy mózgowo związane z ludzką wolą i zamierzonym sprawstwem, a neuronaukowe i filozoficzne interpretacje tych badań wywołały ogromne zainteresowanie. Rezultaty osiągnięte po raz pierwszy przez Libeta sugerują, że mózg podejmuje decyzje o działaniu nieświadomie, zanim działająca osoba zdaje sobie z tego sprawę. Działający podmiot może co najwyżej, w ograniczonym przedziale czasu, świadomie zatrzymać rozpoczęty proces (Libet, 1985). To z kolei prowadzi wielu myślicieli – takich jak amerykański neuropsycholog Daniel Wegner, filozof Derk Pereboom, belgijski filozof Jan Verplaetse czy polski neuronaukowiec Włodzisław Duch – do wniosku, że wolna wola jest niczym więcej niż iluzją.

Zanegowanie wolnej woli, mimo iż w sposób oczywisty wpływa na rozumienie odpowiedzialności moralnej, nie zawsze prowadzi do jej wyeliminowania (Fischer, 2007). Niektórzy teoretycy oddzielają oba pojęcia na tyle ściśle, że możliwe staje się mówienie o odpowiedzialności moralnej przy jednoczesnej całkowitej negacji istnienia wolnej woli (Duch, 2010). Kontrowersyjna możliwość istnienia wolnej woli bez implikacji w postaci odpowiedzialności moralnej również bywa rozważana (Waller, 1998).

Sam Libet, pomimo rezultatów swojego eksperymentu, optował za „filozoficznie realną odpowiedzialnością i wolną wolą” (Libet, 1985, s. 538). Często zauważanymi w tym kontekście problemami – obok wątpliwej możliwości świadomego weta – są założenia Libeta dotyczące natury przyczynowości, tego, czym właściwie miałby charakteryzować się świadomy i wolny wybór, oraz tego, czy badane akty w ogóle przypominają te, za które mielibyśmy być moralnie odpowiedzialni. Pozostaje jednak faktem, że empiryczne dane zdawały się przeczyć wewnętrznemu doświadczeniu sprawczości (Bremer, 2013, s. 65), na którym zarówno ludowa, prawna, jak i filozoficzna myśl nowożytna buduje ideę odpowiedzialności moralnej.

Również w etyce chrześcijańskiej możliwość świadomego i wolnego wyboru traktowana była w sposób oczywisty. Naturalne podstawy działania, w wyniku których postawa lub czyn rzekomo (częściowo lub całkowicie) traci swój woliowy i wolny charakter, podnoszone były głównie w kontekście emocji oraz (prawdziwych lub błędnie ocenianych) zaburzeń psychicznych (Prümmer, 1956, s. 14–19). Co ciekawe, świadomość

ta nie dotyczyła jedynie wpływu negatywnego. Tradycja tomistyczna w pewnym sensie uznaje konstytutywną rolę afektów i emocji w działaniu moralnym (Stump, 2011), jednak wpływ nowożytnego racjonalizmu *de facto* zepchnął nieracjonalne elementy ludzkiego działania poza obszar odpowiedzialności moralnej.

Jednocześnie w tej samej tradycji filozoficzno-teologicznej znajdujemy odniesienia do Arystotelesowskiej teorii cnót, która – wraz z późniejszymi modyfikacjami – na trwałe zakorzeniła się w zachodniej terminologii etycznej i teologiczno-moralnej. Ciągłość ta pozostaje faktem, mimo że do niedawna wpływ tej części arystotelesowsko-tomistycznej spuścizny w etyce ograniczał się głównie do moralnej retoryki oraz do obrazowego ukazywania ideałów moralnych. Pierwotnie pojęcie cnoty (*arête*) bazuje na idei ucieleśnionych dyspozycji mogących znacząco i w stabilny sposób wpływać na ludzkie czyny. Pomimo tego, że rola rozumu i racjonalności (*notabene* pojmowanej zgoła inaczej niż współcześnie) w procesie formowania i realizowania cnoty pozostaje nie do podważenia, podmiotem trwałych dyspozycji do działania (cnoty moralne) pozostaje człowiek jako wegetatywno-zmysłowo-rozumna całość (Arystoteles, 2007).

Upatrywanie źródeł ludzkiego działania, które mogłoby być przedmiotem odpowiedzialności moralnej, wewnątrz podmiotu działającego stało się podstawą chrześcijańskiej i ogólnie zachodniej antropologii. W tym klasycznym ujęciu wola wyraża wewnętrzną zasadę ruchu – czyli *de facto* coś, co ludzie dzielą z innymi zwierzętami. Założenie to, które można nazwać warunkiem *endogeniczności*, podawane jest – obok świadomości celu oraz możliwości wolnego wyboru – jako jeden z warunków zaistnienia *actus voluntaris*, który może być przedmiotem odpowiedzialności moralnej jednostki (Tomasz z Akwinu, 2006, *S.Th.* I-II, 6.1-2). Różnica polega „jedynie” na odmiennych sposobach, przy pomocy których ludzie i zwierzęta (roz)poznają cel własnego działania. Podczas gdy zwierzęta posiadają pewien (zmysłowy) rodzaj poznania, człowiek przewyższa go zarówno stopniem, jak i jakością ze względu na możliwość poznania rozumnego. Mimo to ludzka wola pozostaje umysłowym rodzajem pragnienia, typowym dla gatunku ludzkiego sposobem aktywności w świecie fizycznym.

Postrzeganie woli jako wewnętrznej zasady ruchu ma swoje korzenie w poglądach starożytnych Greków. Oprócz Arystotelesamożna je odnaleźć

np. w myśli stoickiej, do której nawiązywał m.in. św. Augustyn. Według niektórych interpretacji pierwotne użycie przez Augustyna terminu *voluntas* nawiązuje do stoickiej koncepcji *hormê*, która oznacza impuls lub skłonność istoty do działania. Przyczyny dobrowolne (*hormê* lub wewnętrzna zasada ruchu) należą do Boga, aniołów, ludzi i (w analogicznym sensie) do zwierząt (Byers, 2006, s. 174). Dopiero późniejsze prace Augustyna akcentują rolę woli jako ostatecznego arbitra, udzielającego zgody albo Bogu, albo ciału (światu). W obu przypadkach oraz bez względu na osłabienie naturalnymi skłonnościami kojarzonymi m.in. z grzechem pierwotnym podmiot działający pozostaje moralnie odpowiedzialny bardziej p o m i m o aniżeli d z i ę k i naturalnym podstawom działania.

Wiele wieków później Kartezjusz wyraźnie utożsamiał wolę z umysłem – *res cogitans*, które mogą ingerować w działanie mechanistycznej i deterministycznej rzeczywistości *res extensa*. Wola jest tożsama z wolnym wyborem, z naszą możliwością zrobienia czegoś lub nie. Wolność woli była dla Kartezjusza rzeczą oczywistą, a używanie jej było jednoznaczne z działaniem dobrowolnym. Możliwość postąpienia w sposób *a* lub *b* – realizowana za pomocą całkowicie wolnego i niematerialnego umysłu – stanowiła według Kartezjusza podstawę odpowiedzialności oraz „najwyższej doskonałości” człowieka (Kartezjusz, 1984, s. 37).

Wiele późniejszych conceptualizacji woli (w tym etycznych i teologicznych) znajdowało się pod coraz większym wpływem wspomnianych wyżej (wczesnych) idei nowożytnych. Niektórzy myśliciele tego okresu, tacy jak Spinoza, Hobbes czy Hume, wyraźnie umieszczali ludzką (wolną) wolę pośrodku uniwersalnych związków przyczynowych natury. Ten prąd myślowy stał się podstawą współczesnego naturalizmu empirycznego, który wkrótce objął także ludzką wolę, umysł i sprawczość. Mimo tego to właśnie libertariański dualizm Kartezjusza, uzupełniony kantowskimi ideami wolności i odpowiedzialności moralnej, wydaje się mieć największy wpływ na nasze rozumienie działań moralnych (Jarmużewski, 2021). Znaczącą rolę w tym procesie odegrały polityczna i ekonomiczna idea wolności jednostki oraz sukces samej nauki i technologii, który doprowadził do rosnącej wiary w samostanowienie człowieka, potęgę ludzkiego rozumu i rzeczywistość wolnego wyboru.

Wpływ ten nie ograniczył się tylko do racjonalistycznych tendencji (po)oświeceniowych. Widoczny również w chrześcijańskim personalizmie nacisk na konkretny akt moralny, moment wyboru, jego

transcendentny i wolny charakter oraz powiązanie wszystkich tych elementów z doświadczeniem samoświadomości nie są przypadkowe. Stanowią one niejako połączenie myśli tomistycznej i elementów teologii Augustyna z typowo nowożytnym podejściem do wolności i odpowiedzialności. W kontekście Polski nie sposób nie wspomnieć tu wkładu szkoły lubelskiej, a w szczególności Karola Wojtyły i Mieczysława Krąpca, w debatę na temat odpowiedzialności w kontekście jej naturalnych i transcendentnych podstaw.

Opierając się na tradycji scholastycznej i personalizmie, Wojtyła dokonuje wyraźnego rozróżnienia pomiędzy tym, co robi osoba, podejmując działanie (czyn, *actus humanus*), a tym, co po prostu dzieje się z osobą (w osobie). Mimo iż oba fenomeny składają się na szeroko rozumianą dynamikę osoby, jedynie w pierwszym przypadku możemy mówić o działaniu w sensie ścisłym (*actus voluntaris*), czyli o działaniu, za które jesteśmy odpowiedzialni. Według Wojtyły ludzkie działanie moralne – każdy czyn moralny – niesie ze sobą doświadczenie transcendencji (Wojtyła, 1969, s. 30–31, 36, 61–62). Doświadczenie to, jako nieodłączny element aktu moralnego, nie umniejsza jednak leżącego u jego podstaw dynamizmu działającej osoby. Zakładając jedność duszy i ciała w osobie ludzkiej, należy stwierdzić, że na dynamizm ten składają się w dużej mierze elementy neurobiologiczne.

Ujęcie problemowe pojęcia

Choć kwestię odpowiedzialności moralnej w kontekście naturalnych podstaw działania poruszano od starożytności, epoka nowożytna przyniosła nowe wyzwania, które znacząco zmieniły zarówno perspektywę, jak i ton prowadzonej dyskusji. Obejmuje to z jednej strony wspomniany dualizm i racjonalizm, a z drugiej rozwój nauk empirycznych, zwłaszcza psychologii empirycznej w XVIII i XIX wieku, który radykalnie postawił kwestię odpowiedzialności w świetle biofizycznej konstytucji człowieka.

Wielu dziewiętnastowiecznych „ojców psychologii” posiadało szeroką wiedzę z zakresu m.in. fizjologii, a niektórzy – jak William James – twierdzili wręcz, że naukowe badania psychologii powinny opierać się na zrozumieniu systemów biologicznych. W odróżnieniu od tego Freudowska psychoanaliza widziała źródło ludzkich działań głównie

w podświadomości, podczas gdy późniejszy behawioryzm stawiał na relacje bodziec-reakcja, co znacznie zmieniło podejście do zagadnienia wolnej woli i łączonej z nim odpowiedzialności moralnej. Ogólnie rzecz ujmując, spuścizna wczesnych psychologów, takich jak William Carpenter czy wspomniany już William James, wprowadziła zasadniczo nowy sposób myślenia o odpowiedzialności moralnej osoby.

Choć spojrzenie na kwestię uzasadnienia indywidualnej wolności i odpowiedzialności jest w tym przypadku nadal dość konserwatywne, stworzyło ono podstawę do empirycznego uzasadnienia obecnego od wieków sceptycyzmu wobec wolności działania (w ramach klasycznej filozofii i teologii). Paradygmat psychologiczny bez wątpienia przyczynił się też do pojawienia się „iluzjonistycznych” poglądów na kwestię wolnej woli i odpowiedzialności moralnej, typowych dla drugiej połowy XX wieku, a inspirowanych kilkoma dziesięcioleciemiami prac m.in. neuro nauk i neuropsychologii poznawczej.

Podejście zaproponowane w tym kontekście przez Libeta doczekało się wieloaspektowej krytyki. Po pierwsze, zmierzony – czy raczej wyliczony – na podstawie zapisu EEG potencjał gotowości nie musi wyczerpywać wszystkich możliwych przyczyn danego działania. Po drugie, chociaż decyzja o prostym ruchu palcem w niczym nie przypomina tych, za które moglibyśmy być moralnie odpowiedzialni, nawet taki rodzaj wolitywnego aktu jest neurologicznie złożonym i skomplikowanym procesem. Po trzecie, następstwo czasowe pomiędzy aktywnością mózgu a subiektywną (opartą na introspekcji) świadomością intencji niekoniecznie pociąga za sobą następstwo przyczynowe (Bremer, 2013, s. 60–61, 65).

Mimo to niektórzy neurofilozofowie, powołując się na rzekomo udowodnione nieistnienie wolnej woli, sugerują także odrzucenie powiązanego z nim pojęcia odpowiedzialności moralnej. Stanowisko to często kojarzone jest z twardym determinizmem, w przeciwieństwie do różnych form libertarianizmu broniących nowożytnego pojęcia wolnej woli i oczywistości bycia odpowiedzialnym. Obie pozycje łączy przekonanie, że wolnej woli nie da się pogodzić z determinizmem (inkompatybilizm). Semi-kompatybilizm nie odrzuca istnienia odpowiedzialności, jednak nawet stanowiska kom pat y b i l i z m u często podają w wątpliwość powszechne poczucie wolności i autonomii – wraz z konsekwencjami w postaci pewnego zawężenia odpowiedzialności

moralnej podmiotu. Zgodnie z tzw. standardowym argumentem przeciwko wolnej woli nie tylko determinizm, ale także wspomniany wcześniej stochastyzm (przypadkowość zdarzeń) stanowi tu problem. Wyjście z tego impasu może prowadzić poprzez alternatywne paradygmaty w ramach neuronauk oraz przez współczesną biologię ewolucyjną i systemową.

W ostatnich dekadach przedmiotem wielu programów badawczych stało się między innymi ujęcie szerszych podstaw ludzkiego działania. Mimo wielkiej różnorodności, elementem łączącym te nowe podejścia jest wyjście poza analizę zatowizowanych neuronalnych sygnałów w kontekście indywidualnych (i z reguły nieskomplikowanych) aktów motorycznych (rzekomo odpowiadających intencjonalnym stanom mentalnym) na korzyść złożoności i mnogości przyczyn prowadzących do intencjonalnego działania. Jeden z wiodących badaczy na tym polu, Patrick Haggard, trafnie zauważa, że z wolnością aktu wolitywnego wiąże się zakres i złożoność informacji, które – gdy zostają zintegrowane – wywołują akcję (Haggard, 2019, s. 12).

Według Haggarda neuronalną inicjację działania należy rozumieć jako stopniowy zwrot od początkowego stanu płynności, w którym liczne bodźce środowiskowe i wewnętrzne wywołują konkurujące ze sobą skutki, do stanu pewności na krótko przed samym działaniem. Obszar pierwszorzędnej kory ruchowej jest jedynie miejscem, w którym zbiegają się wszystkie etapy wstępne: zachodzące w różnym czasie i w różnych lokalizacjach kory ruchowej (również tych badanych m.in. przez Libeta), niektórych obszarów płatów ciemieniowych oraz, co istotne, odpowiedzialnej za pamięć i świadome intencje kory przedczołowej (Haggard, 2019).

Nowe paradygmaty neuronaukowe przywiązują też należną wagę do długoterminowego aspektu wolnej woli. Niespełna dziesięć lat temu Peter Tse opracował teorię kryterialnej przyczynowości neuronalnej (*critical neural causation*) na podstawie obserwacji, że fizyczne warunki powstawania potencjałów czynnościowych w synapsach podlegają ciągłemu resetowaniu. Według Tsego świadome procesy myślowe, odzwierciedlające stan mózgu w czasie rzeczywistym, mogą wpływać na przyszłe zmiany wyżej wymienionych kryteriów. Taki stan rzeczy może być podstawą istnienia „kryterialnego kodu informacyjnego”, poprzez który podmiot wywiera pośredni wpływ na swoje przyszłe działanie. Działanie

takie, będące wynikiem spełnienia przez bodźce synaptyczne określonych neuronalnych kryteriów ustalonych w (bliższej lub dalszej) przeszłości, nie jest więc ani całkowicie przypadkowe, ani do końca zeterminowane (Tse, 2013). Jako że świadomy wpływ na wspomniane kryteria może odbywać się tak w skali milisekund jak lat, teoria Tsego wspiera zarówno moralną odpowiedzialność za moralne dyspozycje (cnoty), jak i za konkretne akty moralne.

Innym przykładem badań uwzględniających długoterminowy charakter wolnego działania są prace zespołu badawczego pod kierownictwem Stefana Bode'a, sugerujące, że dobrowolne decyzje o działaniu podejmowane są podobnie do decyzji percepcyjnych. Oba rodzaje decyzji są niczym innym niż końcowym efektem procesu zwanego „akumulacją dowodów” (*accumulation of evidence*) ostatecznie przekraczającej próg krytyczny. W przypadku braku wystarczającej liczby dowodów/wskazówek w momencie działania, proces odbywa się na podstawie historii wcześniej dokonanych wyborów, preferencji oraz zakumulowanej wiedzy. Według Bode'a wewnętrznie generowana aktywność mózgu może mieć zarówno charakter stochastyczny, jak i częściowo odzwierciedlać wybory dokonane w przeszłości (Bode et al., 2012). Jednocześnie ze względu na wielość elementów składających się na proces akumulacji świadomy wpływ na podejmowane w mózgu decyzje pozostaje częściowy. Wolność i odpowiedzialność podmiotu za tak generowane działanie leży zatem nie tyle w absolutnej wolności składających się na niego przeszłych decyzji, ile w stopniowym procesie moralnej formacji. Jak wiele innych teorii (na przykład tych związanych z pojęciem neuroplastyczności) długoterminowość świadomej sprawczości może być nośnikiem odpowiedzialności jednostki za proces habituacji, choćby poprzez konsekwentną repetycję właściwych czynów (Arystoteles, 2007, 1103b).

Interpretacja generowania przez mózg intencjonalnych działań jako „akumulacji dowodów” wskazuje na kolejną charakterystyczną cechę współczesnych neuronauk. Jest nią promowane w niektórych kręgach odchodzenie od modelu obliczeniowego (*computational neuroscience*) na rzecz modelu dynamicznego (*dynamical neuroscience*), wykorzystującego formalizm matematyczny (*dynamical systems*) do analizy globalnej aktywności mózgu (na przykład na podstawie zapisów EEG) w czasie. Na znaczeniu zyskują zmienne takie jak tempo,

czas trwania czy dynamiczny przepływ zmian, podczas gdy koncepcja neuronowych reprezentacji stanów psychicznych oraz ich przestrzenna lokalizacja są zwykle odrzucane (Shapiro, 2013, s. 358–359). Stochastyczne elementy tak analizowanej dynamiki mogłyby funkcjonować jako nośnik niezdeteminowanych alternatyw działania (jeden z filozoficznych warunków wolnej woli), podczas gdy elementy porządku (parametry kontrolne) mogłyby umożliwiać woliwną kontrolę, przypisywalność czynu i moralną odpowiedzialność podmiotu.

Aaron Schurger wraz ze współpracownikami argumentuje, że intencjonalne działania nie powstają w żadnym konkretnym miejscu ani nawet w żadnym konkretnym procesie (procesach), ale niejako „wyłączają się z sieci przyczynowej mózgu” (Schurger et al., 2015, s. 761). Według modelu *Stochastic Decision Model* (SDM) to, co nazywamy spontanicznymi aktami woliwnymi, jest wynikiem integracji przypadkowych fluktuacji i bodźców zewnętrznych do poziomu, którego przekroczenie skutkuje reakcją motoryczną. Mierzony przez Libeta „potencjał gotowości” (*Readiness Potential*, RP) nie odzwierciedla procesów prowadzących do decyzji, ale reprezentuje ogólną aktywność mózgu wywołaną przez wiele czynników, co odpowiada rosnącej skłonności podmiotu do działania, a nie jego ostatecznej determinacji (Schurger et al., 2015). Stochastyczne modele decyzji o działaniu sugerują ponadto, że samoświadomość nie jest niezbędnym elementem kontroli działania, ani też faktycznej decyzji. Osobowa kontrola działania zostaje utożsamiona z interakcją wielu procesów, które jedynie w swej całości mogą być traktowane jako neurobiologiczna podstawa odpowiedzialności.

Tak rozumiana wolność wyboru i działania opiera się na naturalnej autonomii charakteryzującej – w różnym stopniu i różny sposób – wszystkie organizmy żywe. Teoria naturalnej autonomii – nazywanej też autonomią biologiczną – wyjaśnia te podstawowe cechy organizmów w kategoriach samodzielnych reżimów autopojezy lub tzw. domknięcia przyczynowego, ang. *causal closure* (Moreno & Mossio, 2015). Rozwinięta w ciągu ewolucji „naturalna autonomia” tłumaczy, w granicach klasycznej biofizyki (bez niepewnych referencji do indeterminizmu fizyki kwantowej), wystarczającą swobodę (dostępne alternatywy działania) i celowość (przypisywalność) działań na poziomie odpowiednim dla poszczególnych gatunków zwierząt. Mimo że trudno w tym przypadku mówić o jakiegokolwiek odpowiedzialności moralnej, trudno też

wyobrazić sobie przypisywalność moralnych aktów u człowieka bez tego rodzaju ewolucyjnego dziedzictwa.

Na naturalną autonomię organizmów składają się kolejno: organizacja i regulacja stymulująca wzrost, utrzymywanie i samonakładanie ograniczeń na własne warunki brzegowe (*boundary conditions*), co umożliwia funkcjonalność i wreszcie relatywnie samodzielną sprawczość. Stopniowy wzrost, zarówno behawioralnej elastyczności, jak i indywidualnej/gatunkowej celowości, może być traktowany jako prekursor specyficznie ludzkiego działania, zakorzenionego w szerszej autonomii ludzkiego organizmu. Racjonalnie wzmocniona i uzasadniona autonomia opiera się więc na ewolucyjnie starszej „wolności od” bezpośrednich przyczyn fizycznych oraz od linearnych relacji przyczynowo skutkowych (Moreno & Mossio, 2015).

Ludzkie zdolności poznawcze uzupełniają te obecne na poziomie somatycznym i behawioralnym. Umożliwiają działającemu człowiekowi efektywniejsze rozważanie dostępnych alternatyw oraz zwiększanie podmiotowości i przypisywalności działania dzięki pośrednim lub długoterminowym mechanizmom (świadomej) kontroli. Tak rozumiane połączenie naturalnej skłonności do eksploracji ze zdolnością do adaptacji i stabilizacji staje się argumentem za istnieniem naturalnie ukształtowanej odpowiedzialności podmiotu (Waller, 1998). Dla pojęcia odpowiedzialności moralnej może to oznaczać głębszą integrację i włączenie somatycznych, afektywnych i ekologicznych aspektów sprawczości, co stanowi jeden z elementów odradzającej się (chrześcijańskiej) etyki cnót (Stump, 2011).

Refleksja systematyczna z wnioskami i rekomendacjami

Poparta badaniami neurobiologicznymi złożoność związków przyczynowo-skutkowych stojących za ludzkim działaniem oraz niebezpośrednie, względne i – w dużej mierze – długoterminowe rozumienie sprawczości i wolnej woli sugerują szersze spojrzenie na kwestię odpowiedzialności moralnej. Skoro neurobiologiczne podstawy działania moralnego z jednej strony wykraczają poza potencjalną rolę samoświadomości w podejmowaniu konkretnych decyzji, a z drugiej uykają prostym

wyjaśnieniom na poziomie biofizycznym czy biochemicznym, to samo dotyczyć może naturalnych podstaw do przypisywania takiemu działaniu osobowej odpowiedzialności.

Wnioski na temat neurobiologicznych podstaw odpowiedzialności moralnej bywają często stosowane w kontekście prawnym i karnym. Ma to na celu wierne odzwierciedlenie problematycznych aspektów odpowiedzialności karnej, szczególnie w sytuacjach nietypowych. Jednocześnie samo pojęcie odpowiedzialności prawnej ma charakter bardzo praktyczny i binarny – podmiot co do zasady albo jest odpowiedzialny (*accountable*), albo nie. Społeczne, psychologiczne i neurobiologiczne okoliczności łagodzące są w tym kontekście postrzegane jako odstępstwa od normalnego stanu osoby, nie jako nieregularności nieodłącznie związane z ludzką konstytucją. To samo podejście daje się zauważyć w odniesieniu do odpowiedzialności moralnej.

Mimo odmiennego znaczenia odpowiedzialność moralna osoby (*responsibility*) jest więc umieszczana mniej więcej na tym samym poziomie, co odpowiedzialność za jej konkretne czyny (*accountability*). W sposób oczywisty obejmuje to wszelkie świadome działania moralne (lub przynajmniej ich końcowy obserwowalny rezultat). Co ważne, taka odpowiedzialność za poszczególne czyny, w kontekście ich neurobiologicznych przyczyn, posiada wspomniane cechy naturalnej sprawczości: pośredniość, długoterminowość, modalność i ekologiczne (społeczne) uwarunkowanie.

Również długoterminowy charakter odpowiedzialności moralnej ma wpływ na kojarzony z konkretnymi czynami aspekt proksymalny. Nawet jeśli w momencie wyboru działanie jest w znacznej części zdeterminowane, doświadczenie uważnego podejmowania decyzji (tak, jak gdyby były one rzeczywiście absolutnie wolne) istotnie wpływa na neurobiologiczno-psychologiczne podstawy przyszłych czynów (Tse, 2013). Z tak rozumianej odpowiedzialności wynika też odpowiedzialność za kształtowanie zamierzeń i planów, ustalanie celów, za integrację i waloryzację konkretnych procesów neuropsychologicznych (takich jak racjonalne myślenie czy kierowanie uwagi), a także za formację oraz działanie oparte na cnotliwych dyspozycjach charakteru (Jarmużewski, 2021).

Przedmiotem odpowiedzialności w tym kontekście mogą być zarówno konkretne ścieżki wyładowań neuronowych, jak również globalne zmiany aktywności mózgu postrzeganego jako układ dynamiczny (Shapiro,

2013). W tym drugim przypadku natura osobowego wkładu woli pozostaje zagadką, zwłaszcza jeśli wyjaśnienia neurokognitywne opierają się na przypadkowych fluktuacjach aktywności neuronowej (Schurger et al., 2015). Mimo to odpowiedzialność podmiotu za tak generowane działanie może być wsparta nawet minimalną obecnością we fluktuującej dynamice nieprzypadkowych informacji (*cues*). Informacje te mogłyby reprezentować proces moralnego namysłu lub osobową historię wyborów i formacji cnoty, za który również możemy być odpowiedzialni.

Cnota w ujęciu Arystotelesa jest zawsze dyspozycją ukierunkowaną na cel lub dyspozycją „dokonującą wyboru”. Cnotliwe działanie, mimo pewnej wyćwiczonej automatyczności, (*hexis*) pozostaje świadomym i rozważnym wyborem, wynikiem aktywnego używania „racjonalnego pragnienia” (*orexis dianoetike*) i „pragnącego rozumu” (*orektikos nous*) (Zagzebski, 1998, s. 217). Zdolność pragnienia dobra, charakterystyczna dla ludzkiej sfery wolitywnej (opartej na praktycznej mądrości – *phronēsis*), musi również skutkować faktycznym wyborem odpowiedniego działania. Każdy wybór i każde działanie zaczyna się od posiadanego już pragnienia lub celu (*telos*), który dzięki racjonalnej deliberacji może stać się podstawą nie tylko dla środków wybieranych do jego realizacji (*means*), ale także dla przyszłych pragnień (celów) działającej osoby (*ends*). Według Arystotelesa absolutnie wolny „wybór” nie jest więc możliwy, a ludzkie działanie moralne jest zawsze wynikiem konkretnych pragnień i racjonalnego namysłu. Nie przeszkadza to jednak silnej normatywie odpowiedzialności za wyuczone i pielęgnowane cnoty (moralne, jak również te intelektualne), za prawidłowe myślenie i za podążanie za celami zgodnymi z ludzką naturą.

Naturalna autonomia organizmów żywych, w tym człowieka, również wspiera takie rozumienie odpowiedzialności. Od strony normatywnej, szczególna pozycja gatunku ludzkiego w świecie istot żywych przywołuje kwestię odpowiedzialności za sposób funkcjonowania w tymże świecie. W obliczu współczesnego kryzysu ekologicznego ludzka odpowiedzialność za chociażby mitygację zmian klimatycznych jest zwykle przedstawiana w kategoriach troski o naturę jako coś zewnętrznego wobec człowieka i jego możliwości działania w świecie. Takie podejście bywa nie tylko nieskuteczne, ale wydaje się nie współgrać z podstawową charakterystyką ludzkiego działania, jaką jest jego ucieleśniony i, co ważne, ekologicznie ukształtowany i współzależny charakter.

Na tej podstawie można argumentować za rozproszonym, choć w istocie szerszym i etycznie realnym pojęciem odpowiedzialności moralnej, wskazując na wielorakie przyczyny ludzkich czynów moralnych. W gąszczu neurobiologicznych przyczyn i prawidłowości staranne umiejscowienie osobistej i dobrowolnej sprawczości podmiotu nie jest zadaniem łatwym. W świetle obecnego stanu wiedzy, ze względu na wielorakie przyczyny poszczególnych aktów moralnych, wolną wolę oraz intencjonalne działanie osoby należy traktować jako neurobiologicznie, ekologicznie i społecznie uwikłane i zależne. Jednocześnie ewolucyjnie uwarunkowana autonomia istot żywych, wraz z typowo ludzkimi zdolnościami poznawczymi, sugeruje możliwość wywierania realnego wpływu na nasze obecne i (w szczególności) przyszłe czyny, w stopniu uzasadniającym szeroko rozumianą odpowiedzialność moralną.

BIBLIOGRAFIA

- Arystoteles (2007). *Etyka nikomachejska*. Przeł. D. Gromska. Warszawa: Wydawnictwo Naukowe PWN.
- Bode, S. et al. (2012). Predicting Perceptual Decision Biases from Early Brain Activity. *Journal of Neuroscience*, 32(36), 12488–12498.
- Bremer, J. (2013). *Czy wolna wola jest wolna? Kompatybilizm na tle badań interdyscyplinarnych*. Kraków: Wydawnictwo WAM.
- Duch, W. (2010). *Czy jesteśmy automatami? Mózgi, wolna wola i odpowiedzialność*. W: P. Francuz (Red.), *Na ścieżkach neuronauk*. Lublin: Wydawnictwo KUL, 219–264.
- Elias, L.J., & Saucier, D.M. (2006). *Neuropsychology: Clinical and Experimental Foundations*. Boston: Pearson/Allyn & Bacon.
- Fisher, J.M. (2007). *Compatibilism*. W: J. Fisher, R. Kane, D. Pereboom, & M. Vargas (Red.), *Four Views on Free Will*. Oxford: Blackwell, 44–84.
- Haggard, P. (2019). The Neurocognitive Bases of Human Volition. *Annual Review of Psychology*, 70(1), 9–28.
- Jarmużewski, M. (2021). *Neurobiology and the Free Will: Implications for Catholic Theological Ethics*. Doctoral Dissertation. KU Leuven.
- Kartezjusz, R. (1984). *Principles of Philosophy*. Przeł. V.R. Miller. Dordrecht: Reidel.
- Libet, B. (1985). Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *Behavioral and Brain Sciences*, 8(4), 529–539.

- Moreno, A., & Mossio, M. (2015). *Biological Autonomy: A Philosophical and Theoretical Enquiry*. Dordrecht: Springer.
- Prümmer, D.M. (1956). *Handbook of Moral Theology*. J.G. Nolan (Red.). Przeł. Gerald W. Shelton. Cork: The Mercier Press.
- Shapiro, L. (2013). Dynamics and Cognition, *Minds and Machines*, 23(3), 353–375.
- Shurger, A., & Uithol, S. (2015). Nowhere and Everywhere: The Causal Origin of Voluntary Action. *Review of Philosophy and Psychology*, 6(4), 761–778.
- Stump, E. (2011). The Non-Aristotelian Character of Aquinas's Ethics: Aquinas on the Passions. *Faith and Philosophy*, 28(1), 29–43.
- Tomasz z Akwinu (2013, 30 października), *Summa Theologiae*. K. Knight (Red.). Przeł. Fathers of the English Dominican Province. <http://www.newadvent.org/summa/> (dostęp: 03.09.2024).
- Tse, P.U. (2013). *The Neural Basis of Free Will: Criterial Causation*. Cambridge, MA – London: MIT Press.
- Waller, B.N. (1998). *The Natural Selection of Autonomy*. Albany: State University of New York Press.
- Wojtyła, K. (1969). *Osoba i czyn*. Kraków: Polskie Towarzystwo Teologiczne.
- Zagzebski, L.T. (1998), *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Aleksandra Kuzior

Politechnika Śląska

<https://orcid.org/0000-0001-9764-5320>

<https://doi.org/10.35765/slowniki.480>

Technologie kognitywne

Streszczenie

DEFINICJA POJĘCIA: Technologie kognitywne to grupa technologii opartych na algorytmach sztucznej inteligencji, maszynowym i maszynowym uczeniu głębokim oraz sztucznych sieciach neuronowych. Wyposażone są w różnorodne funkcje wzorowane na ludzkich procesach myślowych, dlatego są zdolne do uczenia się i wykonywania zadań, które tradycyjnie wymagają ludzkiej inteligencji. Imitują, a czasem nawet przewyższają ludzkie zdolności percepcyjne, poznawcze i decyzyjne.

ANALIZA HISTORYCZNA POJĘCIA: Źródłem zainteresowania i badań nad sztuczną inteligencją i technologiami kognitywnymi należy szukać w badaniach specjalistów z lat 50. i 60. XX w. Przykładowo wymienić tu można prowadzone prace nad serwomanipulatorem, językami programowania oraz programowaniem heurystycznym.

UJĘCIE PROBLEMOWE POJĘCIA: Pod nazwą „technologie kognitywne” rozumiemy interdyscyplinarne połączenie badań podejmowanych przez psychologów oraz rozwiązań technologicznych, które realnie wpływają na rozszerzenie naszych procesów poznawczych. Technologie kognitywne naśladują funkcje ludzkiego mózgu za pomocą narzędzi przetwarzania języka naturalnego, eksploracji danych i rozpoznawania wzorców. Główne problemy w badaniach nad technologiami kognitywnymi to: budowanie zautomatyzowanych i inteligentnych środowisk, implementacja technologii do ludzkiego ciała. W niniejszym opracowaniu skupimy się głównie na pierwszym kierunku.

REFLEKSJA SYSTEMATYCZNA Z WNIOSKAMI I REKOMENDACJAMI: Rozwój technologii kognitywnych obecnie ukierunkowany jest na

tworzenie sztucznego umysłu robotów humanoidalnych i na tworzenie inteligentnych środowisk. Sztuczny umysł ma być imitacją ludzkiego umysłu, co wymaga implementacji procesów psychicznych związanych z myśleniem i świadomością. Jednak obydwu tym procesom powinny towarzyszyć systematyczna refleksja etyczna oraz nakreślenie granic implementacji technologii kognitywnych, abyśmy mogli żyć w lepszym, bardziej humanitarnym społeczeństwie.

Słowa kluczowe: technologie kognitywne, sztuczna inteligencja, generatywna sztuczna inteligencja, kognitywistyka, sztuczne sieci neuronowe

Definicja pojęcia

Terminy „sztuczna inteligencja” i „technologie kognitywne” często pokrywają się zakresowo, niemniej jednak w niniejszym artykule zaznaczamy, że w przypadku sztucznej inteligencji (AI) kluczowe jest znalezienie skutecznego algorytmu w celu wygenerowania najlepszego ogólnego rozwiązania problemu. Z kolei w technologiach kognitywnych nacisk kładziony jest na podejmowanie najbardziej trafnych decyzji na podstawie okoliczności, a co więcej – na dostarczanie informacji potrzebnych do podjęcia takiej decyzji, a nie na faktyczne jej podejmowanie. Technologie kognitywne to grupa technologii opartych na algorytmach sztucznej inteligencji, maszynowym i głębokim maszynowym uczeniu. Zostały wyposażone w różnorodne funkcje wzorowane na ludzkich procesach myślowych, dlatego są zdolne do wykonywania zadań, które zgodnie z tradycyjnym ujęciem wymagają ludzkiej inteligencji. Imitują, a czasem nawet przewyższają ludzkie zdolności percepcyjne, poznawcze i decyzyjne.

Technologie kognitywne (uczące się, poznawcze) nazywane są często technologiami naśladowującymi niektóre funkcje ludzkiego mózgu za pomocą narzędzi przetwarzania języka naturalnego, eksploracji danych i rozpoznawania wzorców. Uznawane są za podzbiór szerszego obszaru sztucznej inteligencji, który z kolei często uważany jest za podzbiór biomimetyków (Rouse, 2018). Według innych, szerszych definicji technologie kognitywne są zbiorem, w skład którego oprócz sztucznej inteligencji wchodzi także zaawansowana analityka, wysoko-wydajne systemy obliczeniowe i systemy cyberfizyczne (Elia & Margherita, 2021).

Technologie kognitywne opierają się na odkryciach kognitywistyki, często dodatkowo wspomaganych przez biologię, neuronaukę, biotechnologię i nanotechnologię. Wykorzystują bazy danych z bioinformatyki i nanoinformatyki, sztuczną inteligencję i nowoczesne narzędzia komunikacji, zwiększając i rozszerzając poznawcze i decyzyjne możliwości człowieka. Nie odtwarzają jeszcze pełnej złożoności ludzkiej inteligencji, ale są na tyle wydajne, że skutecznie potrafią zmaksymalizować komfort życia (Bainbridge, 2006). Jednocześnie jednak badacze, wykorzystując kreatywność ludzkiego umysłu, nadal pracują nad technologią, która w przyszłości może skutecznie z owym umysłem konkurować

(Lakomy, 2023), prowadząc (co jest wysoce prawdopodobne) do jego stopniowej degradacji (Kuzior, 2017).

Technologie kognitywne uznane zostały za megatrendy rozwoju, co oznacza, że szybko się upowszechniają w globalnym środowisku biznesowym i społecznym. Obecnie znajdują zastosowanie w wielu sferach gospodarki i różnych obszarach ludzkiego życia. Rozprzestrzeniają się we wszystkich dziedzinach działalności człowieka, m.in. w medycynie i opiece zdrowotnej, bankowości, finansach, handlu, marketingu, mediach, edukacji, dziennikarstwie, przemyśle, wojskowości. Ich zastosowanie może umożliwić danej firmie zdobycie przewagi konkurencyjnej na rynku (Lakomy, 2023). Już obecnie rynek technologii kognitywnych szacowany jest na miliardy dolarów. Ewoluujący rynek pracy napędzany rozwojem technologii kognitywnych wymagał będzie odpowiednio zaprogramowanego kształcenia w tym zakresie, łącząc zagadnienia techniczne i informatyczne z naukami społecznymi i humanistycznymi. W kształceniu tym konieczne jest uwzględnianie zagadnień etycznych.

Analiza historyczna pojęcia

Pomysły na budowanie myślących maszyn sięgają starożytności, kiedy to greccy mitolodzy wyobrażali sobie takie urządzenia. Przykładem może być rzeźbiony posąg Galatei z kości słoniowej, który ożył w opowiadaniu Owidiusza o Pigmalionie. W XVIII w. prezentowano mistyfikację znaną jako „Mechaniczny Turek”. Była to rzekomo genialna maszyna, rozgrywająca partię szachów. W rzeczywistości w jej środku siedział szachista kierujący ruchami maszyny. W 1997 firma IBM wyprodukowała komputer Deep Blue który pokonał Garry’ego Kasparowa, ówczesnego mistrza świata w szachach. Od tego czasu świat badań nad sztuczną inteligencją i technologiami kognitywnymi dosłownie eksplodował. Inne znaczące daty to 2005, kiedy robot zbudowany w Stanford wygrał wyścig DARPA Grand Challenge, oraz 2011 – wtedy ówczesny symbol technologii kognitywnych Watson (firma IBM) pokonał dwóch największych mistrzów w teleturnieju Jeopardy (bez połączenia z Internetem). Technologia kognitywna zastosowana została także w serwisie mediów strumieniowych Netflix, co w dużej mierze przyczyniło się do spektakularnego sukcesu rynkowego tej firmy (Rouse, 2018).

Technologie kognitywne najczęściej kojarzone są ze sztuczną inteligencją, na temat której bardzo dużo napisano. Stosunkowo mniej opracowań naukowych dotyczy bezpośrednio technologii kognitywnych. Zarówno sztuczną inteligencję, jak i technologie kognitywne kojarzy się głównie z treściami z filmów *science fiction*. Współczesny człowiek często nie zdaje sobie sprawy z tego, że korzysta (w sposób nieświadomy) z rozwiązań opartych na sztucznej inteligencji (ang. *artificial intelligence*, AI) i technologiach kognitywnych (ang. *cognitive technologies*, CT). Prognozuje się, że w najbliższych latach technologie kognitywne będą miały zasadnicze znaczenie w kreowaniu interakcji człowieka z technologią, w szczególności w obszarach automatyzacji i robotyzacji, uczenia maszynowego i technologii IT (Kuzior & Kwilinski, 2022). Interesujących nas źródeł badań nad sztuczną inteligencją i technologiami kognitywnymi należy szukać w opracowaniach specjalistów z lat 50. i 60. XX w., takich jak: John McCarthy, Marvin Minsky czy Alan Turing. Przykładowo wymienić tu można prowadzone w połowie XX w. prace nad serwomanipulatorem służącym do zbadania możliwości bezpośrednich relacji między komputerem cyfrowym a światem fizycznym. Był on zaprogramowany do samodzielnego wykonywania niektórych funkcji zwykle przypisywanych człowiekowi, takich jak postrzeganie świata czy ustalanie sposobów działania (Ernst, 1962). Doszły do tego prace nad językami programowania, uwzględniającymi problematykę badań nad sztuczną inteligencją i symulacją procesów poznawczych (Newell, 1963). Pojawiło się programowanie heurystyczne służące do rozwiązywania wysoce trudnych problemów (Minsky, 1961), a także projektowanie wspólnych systemów kognitywnych człowiek-maszyna (Woods, 1985).

Heurystyczne techniki programowania przyczyniły się do zwiększenia automatyzacji decyzji i stworzyły środowiska interakcji człowieka (naturalnego systemu poznawczego) ze sztucznymi systemami poznawczymi. Przed technologiami kognitywnymi stanęły nowe wyzwania i pojawiło się pytanie, czy można stworzyć hybrydowy system poznawczy składający się jednocześnie z ludzkich i sztucznych systemów poznawczych. Uznany specjalistą w interdyscyplinarnym projektowaniu hybrydowych systemów poznawczych jest Donald Norman – inżynier elektryk i psycholog poznawczy (współzałożyciel i dyrektor firmy konsultingowej User Experience/Usability).

Poznanie i doskonalenie ludzkiego umysłu ukierunkowało naukowców na prace nad wysoce spersonalizowanym komputerowym systemem baz danych, który stałby się prywatnym doradcą, psychoterapeutą lub przewodnikiem duchowym. Przykładem jest ANNE (Analogies in Natural Emotion). Wykorzystując zasadę analogii, ANNE opiera się na emocjonalnej strukturze umysłu indywidualnego użytkownika, zakorzenionej w kulturze. Nie przekazuje autorytatywnie swoich wypowiedzi, mówiąc, co należy robić, ale komunikuje się z użytkownikiem na zasadzie dialogu opartego na symulowanym, intymnym zrozumieniu uczuć użytkownika i pomaga w podjęciu własnej decyzji w danej sytuacji. ANNE wykorzystuje metody psychologiczne, socjologiczne i informatyczne (Bainbridge, 2006). Z kolei rozwój nanotechnologii umożliwił wykorzystanie mikroprocesorów, które wprowadzone do organizmów żywych, w tym ludzkich, mogą informować o wewnętrznym biologicznym stanie i komunikować się z urządzeniami zewnętrznymi.

Technologie kognitywne oferują tym samym nowe sposoby konceptualizacji (tzn. precyzyjnego zdefiniowania procesu badawczego) i komunikowania się w realnym świecie (Bainbridge, 2006). Na czym polegają owe nowe sposoby komunikacji? Wspomnieć należy chociażby projektowanie systemów sztucznej inteligencji wykorzystujących technologie przetwarzania języka naturalnego (*natural language processing*, NLP). Technologie NLP mają na celu umożliwienie maszynom rozumienia ludzkiego języka i reagowania na niego w sposób bardziej naturalny i intuicyjny.

Analizując strukturę języka naturalnego i sposób, w jaki ludzie używają go do komunikacji, systemy NLP można zaprojektować tak, aby zapewniały dokładniejsze i trafniejsze odpowiedzi na dane wprowadzane przez użytkownika. Dodać należy, że technologie kognitywne wyewoluowały dzięki pojawieniu się internetu (głównie sieci i chmury), a dodatkowo jednym z ich przedstawicieli stał się wspomniany już Watson firmy IBM.

Za pomocą sieci neuronowych technologie kognitywne imitują procesy ludzkiego mózgu. Obecnie mapowanie mózgu jest nie tylko metodą diagnostyki chorób i dysfunkcji, ale także wskazówką dla tworzenia sztucznych sieci neuronowych i neurokomputerów. Znaczącą cechą sieci neuronowej jest jej zdolność do uczenia się na bazie przykładów i możliwość automatycznego uogólniania zdobytej wiedzy. Pierwowzorem

sztucznych sieci neuronowych jest ludzki mózg. Sztuczne sieci neuronowe są jednak jego bardzo uproszczonymi modelami. Budowa i odkrywanie właściwości sztucznych sieci neuronowych nie byłoby możliwe bez podstaw zawartych w badaniach neurofizjologicznych i bionicznych. Pierwszą realnie działającą imitacją sieci neuronowej był perceptron. Zachowuje on zdolność poprawnego działania nawet wówczas, kiedy zostanie uszkodzona część jego elementów (Tadeusiewicz, 1993). Model perceptronu z powodzeniem może być wykorzystywany w zarządzaniu przedsiębiorstwem przemysłowym w celu identyfikacji „wąskich gardeł” w działalności biznesowej i poprawy jej funkcjonowania (Kwilinski & Kuzior, 2020). Inne przykłady wczesnych imitacji sieci neuronowych to: Adaline, Limulus, Avalanche, Cerebellatron, Brain State in a Box, Neocognitron, Adaptive Resonance Theory, Mark III, Neural emulator Processor, Mark IV, Odyssey, Crossbar chip, Anza, Parallon, Anza plus (Tadeusiewicz, 1993). Na dzień dzisiejszy zidentyfikować można około 30 typów modeli sieci neuronowych, klasyfikowanych jako spłotowe, rekurencyjne i generatywne.

Obecnie zagadnienia dotyczące sztucznej inteligencji i technologii kognitywnych mają wpływ na prawie wszystkie domeny życia społecznego i indywidualnego. Należą do nich: zarządzanie, rachunkowość i audyt, finanse i bankowość, ubezpieczenia społeczne, rozwój i wzrost gospodarczy, edukacja, systemy inżynierii poznawczej, etyka, neuroetyka. Obserwujemy wpływ na rynek pracy, marketing, jak również na kreację potrzeb i zachowań człowieka, medycynę i opiekę zdrowotną, przetwarzanie języka naturalnego i tłumaczenie maszynowe oraz inne dziedziny (Kuzior & Kwilinski, 2022).

Ujęcie problemowe pojęcia

Technologie kognitywne opierają się na wielkoskalowej architekturze mózgu, modelach funkcji wyspecjalizowanych jego obszarów zwanych w informatyce neurokognitywnej architekturami kognitywnymi. Odwołując się do ścisłej korelacji procesów mózgowych z procesami umysłowymi, można powiedzieć, że architektury kognitywne są zarazem modelami komputerowymi funkcji ludzkiego umysłu. Spośród nich wiele architektur kognitywnych poświęconych zostało opisowi komunikacji

i interakcji pomiędzy człowiekiem i maszyną (Human-Computer Interfaces), a tym samym dotyczą one praktyk kulturowych. Architektury kognitywne oparte na podejściu logicznym symulują wyższe czynności poznawcze (procesy myślenia i rozumowania). Wyróżnić można trzy grupy architektur kognitywnych:

- architektury symboliczne;
- architektury emergentne,
- architektury hybrydowe (Duch, 2010).

W architekturach kognitywnych funkcje poznawcze wspomagane są przez różnego typu pamięć:

- pamięć rozpoznawczą – identyfikację znanych obiektów i dostrzeżenie odstępstw od tego, co znane,
- pamięć skojarzeniową – automatyczne wyciąganie prostych wniosków,
- pamięć proceduralną – pamięć umiejętności manualnych i działań,
- pamięć semantyczną – interpretację sensu,
- pamięć roboczą – kombinatoryczne łączenie różnych informacji w większą całość.

Jednak mózgi ludzkie i „mózgi” komputerów różnią się od siebie w zasadniczy sposób (Duch, 2010). Różnica między mózgiem a komputerem polega na tym, że jako ludzie posiadający mózg potrafimy samodzielnie podejmować decyzje, a także przechowywać ogromną ilość informacji. Z kolei komputer musi być zaprogramowany do wykonywania funkcji i ma ograniczoną zdolność do przechowywania danych i informacji. Różnica dotyczy także neuronalnego i „krzemowego” podłoża procesów zachodzących w mózgu i komputerze.

Technologie kognitywne zwiększają możliwości ludzkiego umysłu i dlatego wymagają właściwego zrozumienia działania tego umysłu. Tajemnice umysłu stopniowo były odkrywane – i nadal są – przez interdyscyplinarną naukę zwaną kognitywistyką (Bremer, 2016). Badania kognitywistyczne prowadzone są na pograniczu psychologii poznawczej, neurobiologii, filozofii umysłu, sztucznej inteligencji, lingwistyki, logiki, fizyki i innych dyscyplin (Duch, 1998). Pozwalają na lepsze zrozumienie właściwości i funkcji ludzkiego mózgu. Kognitywistyka bada procesy uczenia się, języka, myślenia, percepcji, świadomości, podejmowania decyzji oraz inteligencji, zarówno naturalnej, jak i sztucznej. Kognitywistyka wyłoniła się w 1956 roku z cybernetyki Norberta

Wienera. Zasadniczy wkład w jej rozwój wnieśli ponadto Claude Shannon oraz Noam Chomsky. Od 1975 roku w USA uznaje się kognitywistykę (*cognitive science*) za samodzielną dyscyplinę naukową (Lakomy, 2023). Mimo faktu, że poznajemy ludzki umysł od tysięcy lat, używając do tego wiedzy i narzędzi wypracowanych przez wiele dyscyplin naukowych, w dalszym ciągu tajemnice ludzkiego mózgu i umysłu są nie do końca odkryte. Człowiek dąży do poznania i zrozumienia samego siebie. Jednak ludzki umysł w swej warstwie psychicznej, świadomościowej i ludzki mózg w swej warstwie biologicznej są jednocześnie niedoścignionymi fenomenami inspirującymi do dalszych poszukiwań i odkrywania tajemnic.

Postęp techniczny spowodował nie tylko problemy natury teoretycznej, ale także praktyczno-społecznej, dotyczącej zwykłych użytkowników. Człowiek stał się bezradny wobec niektórych problemów generowanych przez dynamiczny rozwój nowoczesnych technologii i nie zawsze potrafi zapobiegać negatywnym skutkom rozwoju nauki i techniki. Nowoczesne technologie uzależniają człowieka od siebie i mogą w negatywny sposób wpływać na jego psychikę, co może z kolei prowadzić do atrofii uczuć, nudy i bezcelowości ludzkiego istnienia.

Rozwój technologii kognitywnych i sztucznej inteligencji obecnie ukierunkowany jest m.in. na zastosowanie technologii kognitywnych do tworzenia sztucznego umysłu robotów humanoidalnych. Przykładem może być humanoidalny robot Phoenix (Canadian Tire Corporation, CTC) z nogami, rękami i głową, będący mniej więcej wielkości przeciętnego człowieka. Waży 70 kg (155 funtów), mierzy 170 cm (5 stóp i 7 cali) i wykazuje „mobilność całego ciała podobną do ludzkiej”. Robot porusza się z prędkością do 3 mil na godzinę (5 km/h) i może przenosić ładunki o wadze do 55 funtów (25 kg). Ale to, co sprawia, że Phoenix jest naprawdę wyjątkowy, to jego sztuczny „umysł”, który został zaprojektowany tak, aby naśladować „podsystemy znajdujące się w ludzkim mózgu”. Oznacza to, że jest on wyposażony w pamięć oraz zmysły wzroku, słuchu i dotyku. Sztuczny umysł ma być imitacją ludzkiego umysłu. Wymaga to jednak implementacji procesów psychicznych związanych z myśleniem i świadomością, a w szczególności modelowania funkcji poznawczych i twórczych przy zastosowaniu sieciowych środków logicznego i asocjacyjnego przetwarzania informacji oraz neuromorficznych sposobów przetwarzania informacji (Stankevich, 2022).

Zainteresowanie problematyką robotów humanoidalnych, w tym robotów społecznych, w ostatnich latach znacznie wzrasta, można powiedzieć, że wręcz w sposób wykładniczy. Tej tematyce poświęcone są międzynarodowe czasopisma naukowe, np. *International Journal of Social Robotics*, *Computers in Human Behavior*, *International Journal of Human-Computer Interaction*, *International Journal of Humanoid Robotics* itd. Na rynku dostępnych jest wiele robotów humanoidalnych, które z powodzeniem zastępują człowieka w różnych sferach ludzkiej działalności, m.in. w kontaktach z klientem, badaniach naukowych, konferencjach itd. Do najbardziej popularnych robotów humanoidalnych należą: Sophia, Pepper, Optimus, Mika, Apollo, Valkyrie, Alsimo i in. Prowadzone są również dalsze prace nad systemami obliczeniowymi, a także integracją technologii kognitywnych i sztucznej inteligencji z innymi technologiami, m.in. *blockchain*. Ważnym aspektem badań jest społeczne postrzeganie interakcji człowiek-robot i możliwości zaspokajania ludzkich potrzeb przez inteligentne maszyny. W badaniach na 700-osobowej populacji wykazano, że człowiek skłonny jest do skorzystania ze wsparcia inteligentnych robotów między innymi w następujących zakresach: odczytywanie nastrojów człowieka (56,7%); odczytywanie ludzkich myśli (27,6%); odczytywanie niewerbalnych aktów komunikacji (57%); odczytywanie i analiza parametrów biomedycznych (53,7%) wraz z zabezpieczeniem pomocy medycznej (46,9%); reagowanie i wykonywanie werbalnych poleceń człowieka (53,6%); rozpoznawanie i zaspokajanie emocjonalnych potrzeb człowieka (47%); rozpoznawanie i zaspokajanie intymnych (seksualnych) potrzeb człowieka (35,9%); bilansowanie i pilnowanie diety człowieka (36,4%) (Kuzior & Kwilinski, 2022).

Rozwój robotów, w tym inteligentnych robotów humanoidalnych, wiąże się w jakimś sensie z powstaniem problemów etycznych. Pojawiają się pytania o możliwości sprawowania etycznej kontroli nad funkcjonowaniem inteligentnych robotów ze względu na fakt, że rozwój sztucznej inteligencji i technologii kognitywnych sprawił, że współczesne roboty są w stanie działać pod wieloma względami podobnie jak człowiek, mogą wykonywać wiele czynności intelektualnych, przypisywanych do tej pory wyłącznie człowiekowi. Tradycyjnie wymiar etyczny dotyczy człowieka i jego działań moralnych jako podmiotu odpowiedzialności. Ze względu na spektakularny rozwój technologii kognitywnych

i sztucznej inteligencji toczy się dyskusja, czy inteligentnym robotom można również przypisać podmiotowość i przyznać im w ten sposób status podmiotów etycznych. Etyczność podmiotów rozpatrywana jest z punktu widzenia sprawstwa czynnego i intencjonalności. Roboty, jako zdolne do określonych działań instrumentalnych, są z pewnością podmiotami (o ile można tutaj użyć takiego określenia) sprawczymi jedynie w sensie mechanicznym, jednakże zakres instrumentalności robotów wyznaczony jest przez człowieka. O sprawstwie biernym mówilibyśmy, gdyby roboty zaniechały jakichś działań. W przypadku intencjonalności wkraczamy w obszar normatywności, a jest to obszar logicznie i pojęciowo przypisywany wyłącznie człowiekowi. Tak samo roju pszczoł nie karzemy za ukąszenie człowieka. Nie można oczywiście wykluczyć możliwości wyposażenia robotów w określony zasób normatywności. Na dzień dzisiejszy jednak nawet najlepiej zaimplementowane reguły etycznego postępowania nie uczynią z maszyny podmiotu moralnej odpowiedzialności, ponieważ odpowiedzialność wymaga wolności. Jeśli maszyna nie mogła postąpić inaczej, niż określają to reguły, to nie może być ani chwalona, ani ganiona za swoje działanie.

Prace programistów w tym zakresie są jednak rozwijane. Gdyby udało się wyposażyć roboty w wartości etyczne i powiązać sprawstwo z intencjonalnością, wtedy można byłoby przypisać robotom jakiś rodzaj podmiotowości, a tym samym odpowiedzialności. Innymi słowy: Aby podmiotowość w sensie biernym zamienić na podmiotowość w sensie czynnym, trzeba spełnić szereg warunków dodatkowych. Kwestie te są nadal przedmiotem dyskusji (Kuzior, 2021).

Technologie kognitywne opierają się na rozwiązaniach generatywnej sztucznej inteligencji. W tym zakresie za przełomowy należy uznać rok 2022, w którym do powszechnego użytkowania trafił ChatGPT, produkt firmy OpenAI, umożliwiający konwersację w czasie rzeczywistym w języku naturalnym, czyli porozumiewanie się człowieka z maszyną i maszyny z człowiekiem. ChatGPT można uznać za innowację przełomową, pozwalającą na prowadzenie dialogu człowieka ze sztuczną siecią neuronową, wykorzystującą technikę głębokiego maszynowego uczenia (*deep learning*). Zainteresowanie ChatemGPT jest bardzo duże i to przykładowo w tak odległych obszarach jak pisanie wierszy w określonym stylu literackim czy uzyskiwanie przepisów na ciasto drożdżowe. Posiada ona jednak ogromny potencjał rozwojowy. Oprócz ChatuGPT

dostępna jest także technologia WormGPT – złośliwa sztuczna inteligencja, nieposiadająca żadnych ograniczeń etycznych. WormGPT wykorzystywany jest przez cyberprzestępców w działaniach hakerskich i phishingu (Lakomy, 2023).

Refleksja systematyczna z wnioskami i rekomendacjami

Generatywna sztuczna inteligencja (Gen-AI) stała się wszechobecna, między innymi w edukacji. Raport edukacyjny „Shifting Horizons – Transformacyjne trendy zmieniające krajobraz szkolnictwa wyższego” (2023) potwierdza konieczność wprowadzenia regulacji dotyczących korzystania z nowych technologii opartych na generatywnej sztucznej inteligencji. Studenci optują za instytucjonalną zgodą na jej wykorzystanie w procesie edukacyjnym. Obecnie ponad 52% polskich studentów korzysta z takich narzędzi jak ChatGPT czy Bard, wykorzystujących duże modele językowe (LLM). Prawie 63% badanych studentów uważa, że ChatGPT pomaga im szybciej wyszukiwać potrzebne informacje i rozwiązywać problemy naukowe (Raport edukacyjny, 2023). Oczywiście deklaracja, że AI pomaga rozwiązywać problemy, nie oznacza, że student-użytkownik wierzy jej bezgranicznie i nie jest świadomy jej ograniczeń. Poniżej omawiam dwa problemy: odpowiedzialne korzystanie z technologii kognitywnych i kwestie etyczne, związane z regulacjami prawnymi.

Jeśli chodzi o odpowiedzialne korzystanie z technologii kognitywnych, to obecnie zaobserwować można dwa możliwe trendy powstałe za sprawą nowoczesnych technologii, w tym technologii kognitywnych, kształtujących nowy typ społeczeństwa powstającego na bazie społeczeństwa informacyjnego i społeczeństwa wiedzy. Pozytywny scenariusz opiera się na wykorzystaniu technologii kognitywnych do rozwoju i poprawy jakości życia człowieka. Negatywnym scenariuszem rozwoju nowego typu społeczeństwa może być polaryzacja społeczeństwa związana z podziałem na twórców i użytkowników branży IT i ICT, które zdominują wszystkie sektory gospodarki i życia ludzkiego. Poza tymi dwoma podmiotami społecznymi znajdują się ci, którzy z powodu analfabetyzmu technologicznego nie będą ani twórcami, ani użytkownikami

nowych technologii. Ów „margines społeczny” nie będzie jednak zagrożony infantyлизacją w takim stopniu, jak użytkownicy bezrefleksyjni, którzy mogą zatracić zdolność samodzielnego myślenia, umiejętność refleksyjnego podejścia do rzeczywistości i naturalną kreatywność. Jeśli lenistwo intelektualne będzie charakteryzować większość użytkowników, może ono doprowadzić do biologicznych i psychologicznych zmian oraz regresu natury ludzkiej i wyróżniających cech gatunku ludzkiego, takich jak: świadomość, wolna wola, moralność, kreatywność, myślenie abstrakcyjne, myślenie symboliczne czy oparta na języku naturalnym umiejętność współpracy. Potrafimy myśleć, rozumować i podejmować decyzje na bazie swoich doświadczeń, motywów i emocji. Dzisiejsze komputery, roboty mogą wykonywać zadania, opierając się na zaprogramowanych instrukcjach, ale nie mogą samodzielnie myśleć ani podejmować decyzji.

Dalszą konsekwencją postępującego zacierania się różnic między człowiekiem a maszyną-robotem wyposażonym w przyszłą, tzn. jeszcze bardziej rozwiniętą sztuczną inteligencję, w połączeniu z lenistwem intelektualnym człowieka może być uzależnienie naturalnej inteligencji od sztucznej inteligencji, którego skutkiem będzie wyuczona bezradność. Wzrośnie także niebezpieczeństwo manipulacji informacją i danymi oraz pogłębi się zacieranie różnic między światem realnym a światem wirtualnym (Kuzior, 2017). Propagowana przez Elona Muska w mediach społecznościowych wizja rozwoju społeczeństwa w koegzystencji z inteligentnymi maszynami jest nieco katastroficzna, jednak niektóre symptomy tego rozwoju są już obecnie zauważalne, jak chociażby samochody bez kierowcy poruszające się w ruchu ulicznym. Kto będzie odpowiedzialny, gdy taki samochód doprowadzi do wypadku drogowego?

Jedną z odpowiedzi na te wyzwania z pewnością są interdyscyplinarne kierunki studiów z zakresu *cognitive technologies*, rozwijane na uczelniach w Holandii (Eindhoven University of Technology), Wielkiej Brytanii (University of Hertfordshire), Niemczech (Universität Osna-brück). Proponowane przez Politechnikę Śląską w Gliwicach studia II stopnia *Cognitive technologies for sustainable development management* są połączeniem tradycyjnej kognitywistyki z technologiami kognitywnymi i stanowią odpowiedź na potrzeby dynamicznie rozwijającego się rynku, który wymaga coraz większej liczby specjalistów, potrafiących

wykorzystywać w swojej pracy nowoczesne technologie, zwłaszcza w obszarze nowoczesnych usług dla biznesu w zakresie zrównoważonego rozwoju.

Wszak to studenci, a nie AI powinni, wykorzystując zdolności swojego umysłu, starać się radzić sobie z rozwiązaniem narastających problemów. AI powinna być jedynie wsparciem. Gotowe podpowiedzi i rozwiązania przez nią podane nie trenują ludzkiego mózgu, raczej przyzwyczajają go do odtwórczego korzystania z gotowych wzorców, które mogą zawierać błędy, wypaczenia i utrwalone stereotypy. ChatGPT może mieć „halucynacje”, czyli bazując na ogromnej ilości danych, przetwarzać informacje w sposób niewłaściwy, tworząc czasem odpowiedzi mylące lub po prostu błędne. Nie oznacza to jednak, że tego typu rozwiązania kognitywne jak ChatGPT należy odrzucać. Raczej trzeba nauczyć się umiejętnie z nich korzystać i przestrzegać określonych zasad.

Przechodząc do kwestii etycznych i regulacji prawnych, należy zauważyć, że ChatGPT jest jednym z najbardziej rozpoznawalnych i rozpowszechnionych przykładów zastosowań technologii kognitywnych na świecie (Lakomy, 2023), jednak jego stosowanie i wykorzystanie może być obarczone licznymi negatywnymi skutkami, takimi jak: zagrożenie nieuprawnionego udostępnienia danych osobowych, podatność algorytmów ChatuGPT na cyberataki, brak jasnych informacji dotyczących przechowywania przez ChatGPT informacji i danych pochodzących z interakcji dialogicznej z licznymi użytkownikami, niejasna polityka prywatności, kwestie etyczne i możliwe konsekwencje prawne związane z prawami autorskimi i własnością intelektualną przy trenowaniu generatywnej sztucznej inteligencji oraz użytkowaniu w ten sposób wytrenowanych algorytmów, nieumyślne generowanie przez ChatGPT dezinformacji i manipulowania treścią, które może np. wpłynąć na wyniki wyborów, podważyć reputację wybranych (czasem losowo) osób lub instytucji, wpłynąć na rynek akcji itd. Wielu użytkowników może paść ofiarą dezinformacji i w sposób nieświadomy przy dobrych intencjach rozpowszechniać fałszywe treści. GenAI może być także użyta do generowania fałszywych danych, służących do potwierdzenia postawionych hipotez naukowych (Raport edukacyjny, 2023).

Ważne z punktu widzenia właściwego wykorzystania oraz rozwoju sztucznej inteligencji i technologii kognitywnych są przepisy prawne i regulacje etyczne. W Unii Europejskiej podjęto prace nad ramami

funkcjonowania „etycznej sztucznej inteligencji”. Prace grupy ekspertów rozpoczęły się w 2018 r. Celem głównym jest uczynienie UE „światowym liderem w dziedzinie innowacji w gospodarce opartej na danych i jej zastosowaniach”. Tak zwana „Biała księga w sprawie sztucznej inteligencji. Europejskie podejście do doskonałości i zaufania” precyzuje, że docelowo powinien być stworzony „ekosystem doskonałości” (ramy polityczne, partnerstwa publiczno-prywatne) oraz „ekosystem zaufania” (ramy regulacyjne, dotyczące ochrony praw podstawowych i praw konsumentów; zachęcenie obywateli do stosowania sztucznej inteligencji). W Księdze znajdujemy siedem zasad niezbędnych do zbudowania godnej zaufania sztucznej inteligencji:

1. przewodnia i nadzorczą rolę człowieka,
2. stabilność i bezpieczeństwo,
3. ochrona prywatności i danych,
4. przejrzystość,
5. różnorodność, niedyskryminacja i sprawiedliwość,
6. dobrostan społeczny i środowiskowy,
7. odpowiedzialność.

Dokument wskazuje także na szanse i zagrożenia związane z rozwojem sztucznej inteligencji. Do szans zalicza się: poprawę opieki zdrowotnej (np. bardziej precyzyjną diagnostykę, lepsze zapobieganie chorobom), zwiększenie wydajności rolnictwa, przyczynienie się do adaptacji do zmian klimatu i ich łagodzenia, poprawę wydajności systemów produkcji, zwiększenie bezpieczeństwa. Do zagrożeń zalicza się: nieprzejrzyste podejmowanie decyzji, dyskryminację ze względu na płeć lub inne czynniki, ingerencję w nasze życie prywatne czy wykorzystanie w celach przestępczych. Regulacje dotyczące etycznej strony rozwoju sztucznej inteligencji muszą odnosić się do respektowania praw podstawowych, takich jak:

1. szacunek dla godności ludzkiej (Ludzi należy traktować jako podmioty moralne, a nie jako przedmioty manipulacji, rankingowania lub sortowania),
2. ochrona wolności jednostki (suwerenności jej decyzji i wyborów, równa dla wszystkich dostępność możliwości, ochrona prywatności i zapewnienie bezpieczeństwa),
3. szacunek dla demokracji, sprawiedliwości i praworządności (AI jako strażnik demokracji, a nie instrument jej zniszczenia i narzędzie manipulacji),

4. zapewnienie równości, niedyskryminacji i solidarności (kluczowe dla uniknięcia niesprawiedliwie stronniczych wyników oraz dbanie o jakość i przejrzystość danych),
5. ochrona praw obywateli (takie same prawa dla wszystkich ludzi).

Na podstawie praw podstawowych sprecyzowane zostały cztery zasady etyczne:

1. szacunek dla ludzkiej autonomii,
2. zapobieganie krzywdom,
3. uczciwość,
4. wiarygodność (Kuzior, 2021).

W Unii Europejskiej w drodze wieloletnich prac i negocjacji powstają regulacje prawne co do wykorzystania sztucznej inteligencji. Ze względu na zakres AI obejmujący także GPAI (duże generatywne modele AI, np. GPT-4 czy Gemini) oraz wysokowydajne systemy regulacje te można uznać za wiążące dla technologii kognitywnych opartych na algorytmach generatywnej sztucznej inteligencji, głębokim maszynowym uczeniu i sztucznych sieciach neuronowych. Uchwalony przez Parlament Europejski akt w sprawie sztucznej inteligencji (Act AI (2024)) proponuje podejście oparte na ryzyku, wyszczególniając przy tym cztery kategorie ryzyka:

- 1) ryzyko minimalne – dotyczy wszystkich modeli AI, których nie zalicza się do trzech pozostałych poziomów ryzyka,
- 2) wysokie ryzyko – systemy AI wywierające niekorzystny wpływ na bezpieczeństwo osób lub przestrzeganie praw podstawowych UE,
- 3) ryzyko bezpieczeństwa produktów objętych sektorowym prawodawstwem unijnym,
- 4) niedopuszczalne ryzyko – szkodliwe zastosowania AI, naruszające prawodawstwo i podstawowe wartości Unii Europejskiej, w związku z czym zakazane będzie stosowanie AI w zakresie:
 - a) oceny społecznej do celów publicznych i prywatnych;
 - b) wykorzystywania osób podatnych na zagrożenia i stosowania technik podprogowych;
 - c) zdalnej identyfikacji biometrycznej w czasie rzeczywistym w miejscach publicznych, wykorzystywanej przez organy ścigania (poza nielicznymi wyjątkami, do których należy 16 rodzajów przestępstw: terroryzm; handel ludźmi; wykorzystywanie seksualne dzieci i tworzenie pornografii dziecięcej; nielegalny handel substancjami odurzającymi i psychoaktywnymi; nielegalny handel bronią, amunicją i materiałami wybuchowymi; morderstwo; ciężkie uszkodzenie ciała; nielegalny handel organami i tkankami ludzkimi; nielegalny handel materiałami jądrowymi lub radioaktywnymi;

- porwania, nielegalne pozbawienie wolności i branie zakładników; przestępstwa podlegające jurysdykcji Międzynarodowego Trybunału Karnego; bezprawne zajęcie samolotu/statku; gwałt; przestępstwa przeciwko środowisku; przestępczość zorganizowana i rozbój z bronią w rękę; sabotaż, udział w organizacji przestępczej);
- d) biometrycznej kategoryzacji ludzi w celu ustalenia rasy, wiary, poglądów politycznych, przynależności do związków zawodowych, orientacji seksualnej;
 - e) indywidualnego systemu nadzoru predykcyjnego;
 - f) rozpoznawania emocji w miejscu pracy i w instytucjach edukacji poza przypadkami podyktowanymi względami medycznymi lub względami bezpieczeństwa;
 - g) pozyskiwania wizerunków twarzy z internetu lub kamer monitoringu w celu tworzenia lub rozszerzania baz danych (Act AI, 2024).

Ze względu na dynamicznie rozwijający się rynek technologii kognitywnych i ich coraz bardziej powszechne użycie istnieje potrzeba kształcenia specjalistów, którzy umiejętnie łączyć będą kompetencje inżynierskie, techniczne z kompetencjami humanistycznymi, etycznymi. Tylko tak przygotowani specjaliści mogą w etyczny sposób kształtować środowisko technologii kognitywnych, czyli architektury kognitywne.

BIBLIOGRAFIA

- Act AI (2024). *Co zmieni AI Act w sprawie sztucznej inteligencji?* <https://odo24.pl/blog-post.co-zmieni-unijny-akt-w-sprawie-sztucznej-inteligencji> (dostęp: 29.09.2024).
- Bainbridge, W.S. (2006). Cognitive technologies. W: *Managing Nano-Bio-Info-Cogno Innovations: Converging Technologies in Society* (s. 203–226). Dordrecht: Springer.
- Bremer, J. (Red.) (2016). *Przewodnik po kognitywistyce*. Kraków: Wydawnictwo WAM.
- Duch, W. (1998). Czym jest kognitywistyka, *Kognitywistyka i Media w Edukacji*, 1, 9–50.
- Duch, W. (2009). Architektury kognitywne, czyli jak zbudować sztuczny umysł. *Neurocybernetyka teoretyczna*, 14, 329–361.
- Elia, G., & Margherita, A. (2021). A conceptual framework for the cognitive enterprise: pillars, maturity, value drivers. *Technology Analysis and Strategic Management*. DOI: 10.1080/09537325.2021.1901874.
- Ernst, H.A. (1962). *MH-1, a computer-operated mechanical hand*. AFIPS Proceedings – 1962 Spring Joint Computer Conference, AFIPS, 142886, 39–50. DOI: 10.1145/1460833.1460839.

- Kuzior, A. (2017). Problem bezrobocia technologicznego w perspektywie rozwoju Przemysłu 4.0. *Etyka Biznesu i Zrównoważony Rozwój*, 4, 31–38.
- Kuzior, A. (2021). *Applied ethics*. Lublin: Wydawnictwo Naukowe Tygiel.
- Kuzior, A., & Kwilinski, A. (2022). Cognitive Technologies and Artificial Intelligence in Social Perception. *Management Systems in Production Engineering*, 2, 109–115. DOI: 10.2478/mspe-2022-0014.
- Lakomy, M. (2023). Technologie kognitywne a demokracja. *BELLONA QUART.*, 3, 31–42. DOI: 10.5604/01.3001.0054.1831.
- Minsky, M. (1961). Steps Toward Artificial Intelligence. *Proceedings of the IRE*, 49(1). DOI: 10.1109/JRPROC.1961.287775.
- Newell, A. (1963). Documentation of IPL-V. *Communications of the ACM*, 6(3), 86–89. DOI: 10.1145/366274.366296.
- Rouse, M. (2018). Cognitive technology. <https://www.techopedia.com/definition/32482/cognitive-technology> (dostęp: 12.12.2023).
- Shifting Horizons – Transformacyjne trendy zmieniające krajobraz szkolnictwa wyższego. (2023). <https://coopernicus.pl/wp-content/uploads/2024/01/Raport-Edukacyjny-PL.pdf>.
- Stankevich, L.A. (2022). Cognitive Technologies and Artificial Mind for Humanoid Robots. *Studies in Computational Intelligence*, 1008 SCI, 3–8. DOI: 10.1007/978-3-030-91581-0_1.
- Tadeusiewicz, R. (1993). *Sieci neuronowe*. Warszawa: Akademicka Oficyna Wydawnicza.
- Woods, D.D. (1985). Cognitive Technologies: The Design of Joint Human-Machine. *Cognitive Systems. AI Magazine*, 6(4), 86–92. DOI: 10.1609/aimag.v6i4.511.

Wykaz Autorów

Dariusz Adamek, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0002-5570-2259>

Nina Bażela, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0009-0003-1147-1086>

Jacek Bielas, Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-9771-4162>

Andrzej Bielecki, AGH Akademia Górniczo-Hutnicza

<https://orcid.org/0000-0002-0192-3785>

Józef Bremer, Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-9664-8896>

Andrzej Dąbrowski, Uniwersytet Komisji Edukacji Narodowej
w Krakowie

<https://orcid.org/0000-0002-3998-9608>

Mariusz Flasiński, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0002-0910-4928>

Paweł Fortuna, Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0002-0633-4453>

Mateusz Hohol, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0003-0422-5488>

Ewa Jaglarz, Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0001-8754-8770>

Mateusz Jarmużewski, Windesheim University of Applied Sciences

<https://orcid.org/0000-0002-8398-1398>

Piotr Konderak, Uniwersytet Marii Curie-Skłodowskiej w Lublinie

<https://orcid.org/0000-0002-9161-4870>

Aleksandra Kuzior, Politechnika Śląska

<https://orcid.org/0000-0001-9764-5320>

Marek Lechniak, Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0002-0768-7963>

Ewa Odoj, Uniwersytet Ignatianum w Krakowie

<https://orcid.org/0000-0003-0821-9317>

Anita Pacholik-Żuromska, Uniwersytet Mikołaja Kopernika w Toruniu

<https://orcid.org/0000-0002-8674-9412>

Marek Pokropski, Uniwersytet Warszawski

<https://orcid.org/0000-0002-7423-6951>

Sławomir Sztajer, Uniwersytet im. A. Mickiewicza w Poznaniu

<https://orcid.org/0000-0001-8975-2994>

Michał Wierzchoń, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0002-7347-2696>

Łukasz Wiraszka, Uniwersytet Jagielloński w Krakowie

<https://orcid.org/0000-0002-3653-5926>

Zbigniew Wróblewski, Katolicki Uniwersytet Lubelski Jana Pawła II

<https://orcid.org/0000-0003-4477-6903>

