

Wrażliwość estetyczno-logiczna w badaniu naukowym jako wyzwanie dla Sztucznej Inteligencji

Myśl, że uczony w swoich badaniach dąży do piękna jest w refleksji nad nauką czymś nowym.¹ Według danych, które udało mi się zebrać, po raz pierwszy pojawia się ona u Henri Poincaré'go (1854-1912) w jego pracach z początku tego stulecia. Wcześniej, filozofowie i uczeni, od Platona do Leibniza i dalej, wyrażali swój podziw dla piękna świata, skąd wzięła się też w grece nazwa *kosmos*, oznaczająca ład i ozdobność. Nie mówiło się jednak o pięknie jako takiej własności, że gdy cechuje ona teorię nadaje wiarygodność jej roszczeniom do trafnego poznawczo odzorowania świata.

Pojawienie się pojęcia piękna teorii zbiega się z narodzinami dylematów, które były obce dawniejszym uczonym, a stanęły przed współczesnymi. Są to problemy wyboru pomiędzy teoriami. Zważmy, że nie było ich jeszcze w czasach Kanta. Dla Kanta było bezspornym punktem wyjścia, że mechanika Newtona jest teorią jedyną i ostateczną; pozostaje tylko wytłumaczyć, jak to możliwe, że umysł jest w stanie do niej dojść; co też próbował wyjaśnić swą nauką o wrodzonych kategoriach umysłu.

Problem przeżywania przez uczonych piękna teorii da się traktować w kategoriach psychologicznych. Tym jednak, co rzeczywiście interesujące jest aspekt *epistemologiczny*. Wyraża się on w pytaniu, czy jest poznawczo prawomocne kierować się w wyborze teorii względem na jej piękno, objawiające się w takich cechach jak prostota, albo wewnętrzna spójność, podobna jakiejś harmonii.

Ci, którzy na to pytanie odpowiadają twierdząco, dzielą się na dwa nurty. Jeden z nich traktuje twórczość naukową na wzór twórczości w sztuce, rezy-

¹ Część niniejszego tekstu była referowana – pod tytułem „Prawda a piękno: wyznania wielkich uczonych” – na konferencji metodologicznej zorganizowanej przez Instytut Informatyki UJ, Kraków, 6-8 maja 1998. Całość została opracowana jako wkład do projektu badawczego KBN nr 8T11C01812 (Zespół Automatyki, Elektroniki, Informatyki i Telekomunikacji). Występujące w tytule pojęcie wrażliwości estetyczno-logicznej jest określone w odcinku 4, poprzedzając podjęte w odc. 5 (na podstawie zgromadzonej uprzednio dokumentacji historycznej) zagadnienie konstrukcji lub symulowania tej dyspozycji w SI.

gnując z poszukiwania obiektywnej prawdy, a za cel ostateczny mając jakąś kreację, podobną celom artystycznym. Wyrazistym przedstawicielem tego nurtu był Jan Łukasiewicz, a pewne jego elementy można znaleźć u Poincaré'go.

Inny nurt traktuje piękno na kształt kryterium prawdy. Jest to kryterium nieoficjalne, subiektywne, nie mające mocy intersubiektywnego argumentu w dyskusji naukowej, jednakże mające wpływ na losy nauki. Daje ono badaczowi tę siłę przekonania co do trafności wyboru, bez której nie byłby on w stanie inwestować w żmudne i kosztowne badania prowadzące do znalezienia argumentów intersubiektywnych, czy to eksperymentalnych czy dedukcyjnych. Bywają sytuacje, gdy to, jaki kurs weźmie nauka zależy od decyzji wielkiego uczonego, pioniera nowych idei, a jego decyzja ma za kryterium piękno. Wśród rzeczników takiego myślenia można odnaleźć Einsteina i Heisenberga, a do ich wypowiedzi można dołączyć niektóre sformułowania Poincaré'go, którego poglądy oscylowały także i w tym kierunku.

Każdemu z tych nurtów poświęcimy osobną część rozważań.

1. Piękno jako cel nauki

Zaczynając od Poincaré'go jako inicjatora naszej problematyki, trzeba zwrócić uwagę, że był on głównym rzecznikiem *konwencjonalizmu* w filozofii nauki. Kierunek ten utrzymuje, że teoria naukowa jest swobodną kreacją umysłu, zmierzającą do możliwie najprostszej i najdogodniejszej interpretacji rzeczywistości. Pierwszy człon tego programu – wolność tworzenia – upodobia naukę do sztuki. Pojawia się więc dla tych dziedzin jakby rodzaj nadrzędny, w obrębie którego różnicuje naukę dążenie do tworzenia obrazu rzeczywistości (podczas gdy artysta może poprzestać na ekspresji swych przeżyć).

Konwencjonalizm Poincaré'go nie jest postulowaniem prawa do arbitralności w wyborze teorii. Jeżeli można mówić o opozycji pomiędzy nim a realizmem, to tylko wtedy, gdy realizm w nauce pojmuje się jako program oddawania faktów w sposób niejako fotograficzny, bez dokonywania między nimi wyboru. Wybór zaś to jakaś decyzja, a gdy decyzja zostaje uzgodniona w społeczności uczonych, można ją nazwać umową czyli konwencją.

Tak docieramy do kwestii motywów wyboru i tu się pojawia pojęcie piękna. Sam Poincaré tak o tym pisze (kursywa ode mnie – WM).

«Uczony bada naturę ponieważ się nią zachwyca, a zachwyca się nią, ponieważ jest piękna. Oczywiście, nie mówię tu o pięknie odbieranym przez zmysły; daleki jestem od tego, by go nie doceniać, lecz nie ma ono nic do

czynienia z nauką. Mam na myśli to głębsze piękno, które się bierze z harmonijnego układu elementów i które może uchwycić czysta inteligencja. Dla tego intelektualnego piękna, bardziej niż dla przyszłego dobra ludzkości, uczony oddaje się długim i trudnym pracom. Poszukiwanie więc szczególnego piękna, poczucie harmonii wszechświata jest tym, co nami powoduje, żeby wybierać te fakty, które się najbardziej przyczyniają do owej harmonii; tak jak artysta wybiera z rysów swego modelu te, które czynią obraz doskonałym i nadają mu wyraz i życie. A nie trzeba się obawiać, że to instynktowne pragnienie odwiedzie uczonego od szukania prawdy. Jakże dalece harmonia realnego świata góruje nad harmonią sennych zjaw!» [W domyśle: a więc harmonia świata jest kryterium prawdy obrazu świata, gdy sam ten obraz cechuje się harmonią.] H. Poincaré *The Value of Science* (Dover Publications 1958), 'Prefatory to [English] Translation', s. 8 (przekład *ad hoc* – WM; cytowane ze skrótami).

Problem wyboru faktów naukowych kierowanego odczuciem piękna doznał się ważkiej refleksji w pewnym artykule poety i zarazem teoretyka poznania, jakim był Paul Valéry, artykule powołującym się na myśl Poincaré'go. Oto jego fragment.

«Trzeba dwu rzeczy, aby cokolwiek wynaleźć. Jedna tworzy kombinacje; druga wybiera wśród masy obiektów dostarczonych przez pierwszą; rozpoznaje te, na których jej zależy i które są dla niej ważne. To, co nazywamy geniuszem, jest w o wiele mniejszym stopniu pracą pierwszą, a w o wiele większym — umiejętnością drugą; ocenienie wartości tego, co przed nią leży, i dokonanie odpowiedniego wyboru.» Valéry powiada dalej, że reguły takiego wyboru (tu cytuje Poincaré'go) «są niezmiernie subtelne i misterne, niepodobna niemal wyrazić ich w ścisłym wystąpieniu; czuje się je raczej niż formułuje; jakżeby się wobec tego miało wyobrazić sobie sito zdolne mechanicznie je stosować?» P. Valéry, „De la simulation”, *Nouvelle Revue Française*, vol. 28, Mai 1927, s. 620. Cytowane za książką J. Hadamarda, s. 38 (zob. niżej).

Wypowiedzi Valéry'ego zaczerpnięte są z cytatów zgromadzonych przez wybitnego matematyka J. Hadamarda w jego rozważaniach o twórczości naukowej. Hadamard podsumowuje je w sposób następujący.

«Począwszy od Poincarégo wiemy, że poczucie piękna odgrywa niemałą rolę i jest *niezbędnym składnikiem* samego procesu odkrywania. Tak więc, doszliśmy do dwóch wniosków:

że odkrycie jest wyborem,

że kieruje nim w zdecydowany sposób poczucie naukowego piękna.»

Jacques Hadamard, *An Essay on the Psychology of Invention in the Mathematical Field*, Dover Publications 1954 (wyd. oryginalne 1945). Polski przekład *Psychologia odkryć matematycznych*, tłum. R. Molski, PWN 1964, seria Omega, S. 39.

Podobne myśli znajdujemy w rozprawce Jana Łukasiewicza na temat twórczości w nauce. Łukasiewicz znał odnośne teksty Poincarégo, jak wynika z przypisów, ale w sprawie piękna w nauce nie powołuje się na tego myśliciela (przypisy dotyczą innych tematów). Podobieństwo jest jednak nie do zaprzeczenia, choć Łukasiewicz idzie dalej w pomniejszaniu roli prawdy. Czytając Poincaré'go nie wiele się dowiadujemy o stosunku prawdy i piękna, ale prawda gdzieś widnieje w tle wraz z domniemaniem, że piękno jakoś na nią naprowadza. U Łukasiewicza zaś czytamy, co następuje (wyróżnienia kursywą – cyt. autora).

«Błędne jest przeto zdanie, że celem nauki jest *prawda*. Nie dla prawdy umysł tworzy. *Celem nauki jest budowa syntez zaspokajających ogólnoludzkie potrzeby intelektualne.* [...] Twórczość poetycka nie różni się od naukowej większym polotem fantazji. Kto, jak Kopernik, ziemię ruszył z posad i pchnął ją na tory w krąg słońca, lub, jak Darwin, ujrzał w mgłę dziejów genezyjskie przemiany gatunków, godzien stanąć obok największych poetów. Uczony jednak tym się różni od poety, że zawsze i wszędzie *rozumuje*. Nie wszystko musi i może uzasadnić, ale cokolwiek głosi, musi węzłami logicznymi powiązać w ścisłą całość. Na dnie tej całości leżą sądy o faktach, nad nimi wznosi się teoria, która te fakty tłumaczy, porządkuje, przepowiada. Tak powstaje *poemat nauki*.»
Jan Łukasiewicz, „O twórczości w nauce” w: *Księga pamiątkowa ku uczczeniu 250 rocznicy założenia Uniwersytetu Lwowskiego*, Lwów 1912. Przedruk w: Jan Łukasiewicz, *Z zagadnień logiki i filozofii. Pisma wybrane*, wybór i wstęp Jerzego Śłupeckiego. PWN, Warszawa 1961. Cyt. za tym wyborem, s. 74.

Myśli Łukasiewicza, podobnie jak Poincaré'go, przypadają na okres poprzedzający nową wielką erę fizyki, którą otworzyły teoria względności i teoria kwantów, w której się też mieści odkrycie ewolucji wszechświata. Ta zdumiewająca seria odkryć miała zapładniający wpływ na filozofię nauki, uprawianą intensywnie przez samych uczonych (nie była to bynajmniej praca amatorska „po godzinach” pracy zawodowej, bo refleksja nad nauką, poznaniem i umysłem ściśle się spletała z samymi badaniami). Widać to, na przykład, w autobiografii intelektualnej Heisenberga. Z niej zaczerpniemy pewien epizod pomocny w rozwijaniu następnego punktu.

2. Piękno jako heurystyczne kryterium prawdy

W roku 1926 młody docent Uniwersytetu w Getyndze Werner Heisenberg został zaproszony do Berlina, samej stolicy ówczesnej fizyki, gdzie działali Planck i Einstein, żeby na świeżo utworzonym kolokwium wygłosić odczyt o jego własnych wynikach i ideach w teorii kwantów. Po odczycie Einstein zaprosił młodszego kolegę na spacer dla kontynuowania dyskusji. Heisenberg tak relacjonuje jej końcowy fragment.

«Einstein spojrział na mnie trochę krytycznie.

— „Dlaczego właściwie wierzy pan tak mocno w swą teorię, skoro w gruncie rzeczy tak wiele pytań, i to kluczowych, jest jeszcze całkowicie niewyjaśnionych?”

Odpowiedziałem wtedy coś następującego.

— Wierzę tak samo jak pan, że prostota praw przyrody ma jakiś charakter obiektywny, że chodzi tu nie tylko o ekonomię myślenia. Gdy sama przyroda prowadzi nas ku formom matematycznym wielkiej prostoty i piękności – pod formami rozumiem spójne systemy założeń, aksjomatów itp. – ku formom, których wcześniej nikt nie wymyślił, to trudno wtedy nie uwierzyć, że są one «prawdziwe», to znaczy, że przedstawiają jakiś rzeczywisty rys przyrody. Być może, że te formy mają coś do czynienia także z naszym stosunkiem do przyrody, że istnieje w nich także element ekonomii myślenia. [...] Może mi pan zarzucić, że stosuję tu jakieś estetyczne kryterium prawdy, gdy mówię o prostocie i pięknie. Ale muszę przyznać, że wielka siła przekonania bierze się we mnie z prostoty i piękna tej matematycznej struktury, którą przyroda nam podpowiada. [...] Nie oczekuję bynajmniej, że przez to rozwiążą się wspomniane przez nas wcześniej trudności teoretyczne. *Prostota struktury matematycznej sprawia, że musi być możliwe wymyślenie wielu eksperymentów których wynik da się wyliczyć z wielką dokładnością dzięki danej teorii.* [kursywa – WM] Kiedy przeprowadzi się eksperymenty i przyniosą one ów przewidziany wynik, wtedy nie można już mieć wątpliwości, że nasza teoria trafnie w danym zakresie oddaje stan rzeczy.» Werner Heisenberg, *Der Teil und das Ganze. Gespräche im Umkreis der Atomphysik*, Piper & Co., München 1969. (przekład *ad hoc* – WM), s. 85n.

Dalszy ciąg rozmowy dotyczy kryteriów prawdy w fizyce oraz pytania, czy wszystko da się sprawdzić doświadczalnie. Zacytowany fragment z autocyta-tem Heisenberga rzuca snop światła na stosunek piękna do prawdy (w zdaniu wyróżnionym kursywą). Chodzi o to, że im bardziej teoria jest obdarzona estetycznym walorem prostoty, tym bardziej jest sprawdzalna, daje bowiem większe możliwości zaprojektowania testów eksperymentalnych. Myśl, że teoria sprawdzalna musi być prosta można poprzeć przykładami negatywnymi wskazanymi przez Poppera. Są to przykłady, które świadczą, że im teoria bardziej zawiła, oblepiona „przybudówkami” typu hipotez *ad hoc* oraz furtkami interpretacyjnymi (Popper wymieniał tu przykładowo marksizm i psychoanalizę), tym mniej poddaje się sprawdzeniu.

Dostarcza więc Heisenberg propozycji, jak rozumieć stosunek piękna do prawdy. W wymiarze osobistego przeświadczenia badacza, piękno – utożsamiane z prostotą – może stanowić bezpośredni motyw uznania teorii za prawdziwą. W wymiarze zaś intersubiektywnym cecha ta nie jest bezpośrednim kryterium prawdy, lecz do niej przybliża, jest bowiem pewnym aspektem cechy sprawdzalności.

Dzięki relacji Heisenberga zyskujemy też pogląd na postawę Einsteina w rozważanej kwestii. Myśli o prostocie i harmonii świata odzwierciedlanej w naszych teoriach stanowią wspólną płaszczyznę obu rozmówców. Einstein wypowiadał się wielokrotnie w tym duchu, sławiąc harmonię przyrody. Przykładem może być jego przemówienie na obchodach 60 rocznicy urodzin Maxa Plancka wygłoszone na posiedzeniu Towarzystwa Fizycznego w Berlinie. Tę samą postawę przypisuje on Planckowi, a ponieważ trudno odmówić wiarygodności sądowi Einsteina na temat jego bliskiego przyjaciela, trzeba umieścić również Plancka w szeregu uczonych inspirowanych wizją harmonii świata.

Einstein posługuje się w tym tekście Leibnizjańskim pojęciem odwiecznej harmonii (*harmonia praestabilita*, ang. *pre-established harmony*), które chwali jako trafne określenie stosunku między światem zjawisk i zasadami teorii naukowej. Zauważa przy okazji, że filozofowie, w odróżnieniu od fizyków, za mało zwracają uwagi na tę relację. Powiada on dalej.

«Pragnienie podziwiania tej odwiecznej harmonii jest źródłem tej niewyczerpanej cierpliwości i wytrwałości, z jaką Planck poświęcił się, jak widzimy, najbardziej ogólnym problemom naszej nauki. [...] Stan umysłu, który czyni człowieka zdolnym wykonywać tego rodzaju pracę jest podobny uniesieniu religijnemu albo stanowi zakochania; codzienny trud nie pochodzi z jakiegoś postanowienia czy programu, lecz prosto z serca.»
Albert Einstein, *Essays in Science*, Philosophical Library, New York 1934 (przekład *ad hoc* – WM), s. 5.

Przemówienie kończy się w znamienity sposób, bo życzeniem, żeby jubilatowi udało się połączyć elektrodynamikę i teorię kwantów w jeden logicznie spójny system (*a single logical system*). W bezpośrednim kontekście cytowanych wyżej zdań, jest to kolejny przyczynek do wyjaśnienia, na czym polega owa pożądana prostota i harmonia: im więcej zakresów zjawisk teoria obejmuje, dostarczając jednego wyjaśnienia w miejsce wielu, tym więcej wprowadza prostoty do danej dyscypliny i tym głębiej odśłania harmonię świata.

3. Głosy innych uczonych

W potocznym myśleniu silny jest stereotyp przeciwstawiający naukę sferze piękna artystycznego. Toteż przytoczone wypowiedzi uczonych mogą być odebrane jako te wyjątki, które wedle potocznego powiedzenia „potwierdzają regułę”. Ktoś może by argumentował, że zdarzają się połączenia przeciwstawnych talentów i zamiłowań w jednej osobowości, ale nie należy w tym upatrywać prawidłowości ogólnej.

Aby pokazać, że nie są to głosy nielicznych solistów, posłuchajmy dwóch świadectw mówiących o powszechności kryterium piękna w postępowaniu naukowym. Jedno pochodzi od matematyka, drugie od fizyka.

W autobiografii Stanisława Ulama *Przygody matematyka* (w oryginalnej *Adventures of a Mathematician*, University of California Press, 1976) nie znajdujemy refleksji poświęconych specjalnie temu problemowi (jak te u Poincaré'go czy Heisenberga), są natomiast wzmianki przy okazji innych tematów; sposób czynienia tych wzmianek świadczy, że obecność piękna w dociekaniach matematycznych jest dla autora czymś oczywistym, co nie wymaga argumentacji ani osobnej uwagi.

Rozważając np. jak przebiega praca naukowa w różnych fazach życia, powiada, że podziw dla piękna pozostaje do końca, cytując w tym miejscu zdanie Einsteina:

«Najpiękniejszą rzeczą, jakiej możemy doświadczyć, jest tajemnica. Jest to źródło całej prawdziwej nauki i sztuki.» s. 316 w wersji polskiej, Pruszyński i S-ka, Warszawa 1996, przekład A. Górnickiej.

Na innym miejscu, gdy powiada, że ogół matematyków jest zgodny co do wartości wyników i teorii, tłumaczy to podzielanym przez ten ogół odczuciem piękna.

«W odczuciu piękna, jakie daje matematyka, musi więc być coś obiektywnego, jeśli nawet nie zostało to jeszcze zdefiniowane.» Op. cit., s. 304.

Fizyk Roger Penrose, który wsławiwszy się m.in. wynikami na temat czarnych dziur, stał się również autorem klasycznych już prac o sztucznej inteligencji, w tym drugim kontekście rozważa rolę natchnienia jako procesu nie podlegającego algorytmizacji, a w związku z tym rolę odczucia piękna (podkreślenia kursywą – od Penrose'a).

«Silne przekonanie o *śluszości* idei, która pojawiła się w krótkiej chwili natchnienia [...] jest ściśle związane z jej cechami estetycznymi. Jest znacznie bardziej prawdopodobne, iż piękna idea okaże się prawdziwa, niż że przydarzy się to idei brzydkiej.»

Tu powołuje się Penrose na innych jeszcze autorów oraz cytuje wykorzystany wyżej tekst Valéry'go (odc. 1, ostatni z cytatów z książki Hadamarda). I dalej pisze.

Również Dirac przyznał bez zażenowania, że to właśnie *poczucie piękna* umożliwiło mu odkrycie równania elektronu, podczas gdy inni poszukiwali go na darmo. [...]

Wydaje mi się oczywiste, że kryteria estetyczne odgrywają ważną rolę nie tylko w natychmiastowej ocenie idei, które zawdzięczamy chwilom

natchnienia, ale również w podejmowaniu znacznie częstszych decyzji związanych z codzienną pracą naukową. Ścisły dowód pojawia się zazwyczaj *na końcu!* Przed jego sformułowaniem trzeba wiele rzeczy odgadnąć, w czym ogromnie pomagają przekonania estetyczne, zawsze ograniczone argumentami logicznymi i znanymi faktami.»

Roger Penrose, odc. pt. „Natchnienie, wgląd i oryginalność” w rozdz. 10 książki *Nowy umysł cesarza. O komputerach, umyśle i prawach fizyki*, Wyd. Naukowe PWN, Warszawa 1995. Przekład P. Amsterdamskiego. Oryginał: *The Emperor's New Mind. Concerning Computers, Minds and the Laws of Physics*, Oxford University Press, 1989.

Penrose czyni ważną uwagę dotyczącą sprawy, której nie podejmowali inni autorzy: jaka jest wiarogodność owego natchnienia kierowanego odczuciem piękna? Czyni to w zdaniu „Jest znacznie bardziej prawdopodobne, iż piękna idea okaże się prawdziwa, niż że przydarzy się to idei brzydkiej.” W innym miejscu (opuszczonym przeze mnie przy cytowaniu), dotyczącym jego własnych tego rodzaju przeżyć, powiada, że wyczuł pewne rozwiązanie „w nagłym przeblysku intuicji, ale tylko z sześćdziesięcioprocentową pewnością”.

Jeśli istotnie kryterium piękna w sferze poznawczej nie jest niezawodne, zasługuje ono na większe zaufanie niż gdyby mu przypisać niezawodność absolutną. Jest wtedy bowiem czymś ze świata realnego, a nie świata poddanego jakiejś mitologizacji. Jest fenomenem wymagającym wyjaśnienia, mało dotąd zrozumiałym, ale rokującym szanse wyjaśnienia po odpowiednio wnikliwym badaniu.

Spróbujmy pierwszego przybliżenia, podejmując tę próbę w kontekście pewnych założeń dotyczących inteligencji. Jest to ten sam kontekst, w którym porusza się Penrose, podejmując temat „natchnienia” w *The Emperor's New Mind* — studium poświęconym sztucznej inteligencji i jej stosunkowi do naturalnej.

4. Wrażliwość estetyczno-logiczna a inteligencja

Cechę opisywaną wyżej w cytowanych tekstach będę określał jako *wrażliwość estetyczno-logiczną*, w skrócie *el-wrażliwość*. Sens pierwszego członu w „el” jest oczywisty w świetle cytowanych i komentowanych wyżej tekstów. Człon „l” wskazuje na związek z problematyką metodologiczną, a więc (w szerszym sensie) logiczną, oraz na to, że istotną rolę w funkcjonowaniu tej dyspozycji ma podświadome przewidywanie konsekwencji logicznych.

Dopiszmy el-wrażliwość do listy cech, które stanowią składniki czy też warunki (trudno oddzielić jedne od drugich) wielce złożonej cechy zwanej

inteligencją. Na liście tej znajdują się mniej dyskusyjne cechy, które dziś łatwo wyróżnić dzięki analogiom z komputerami. Należy do nich pojemność pamięci operacyjnej (chwilowej), pojemność pamięci zewnętrznej i zgromadzone w niej zasoby wiedzy, organizacja wiedzy warunkująca sprawność przeszukiwania (zapewnia ją człowiekowi odpowiedni aparat pojęciowy), szybkość przetwarzania informacji itp.

Wszystkie one, w procesie myślenia twórczego, przyczyniają się do wykonywania pierwszej rzeczy spośród tych dwu, które wymienia cytowany Paul Valéry. Przypomnijmy tę ważną wypowiedź (tym razem z wyróżnieniami kursywą dla celów obecnej analizy).

«Trzeba dwu rzeczy, aby cokolwiek wynaleźć. Jedna tworzy *kombinacje*; druga *wybiera* wśród masy obiektów dostarczonych przez pierwszą; rozpoznaje te, na których jej zależy i które są dla niej ważne. To, co nazywamy geniuszem, jest w o wiele mniejszym stopniu pracą pierwszą, a w o wiele większym — umiejętnością drugiej; ocenienie wartości tego, co przed nią leży, i dokonanie odpowiedniego wyboru.»

Pamięć, wiedza, szybkość przetwarzania — to czynniki, które zapewniają odpowiednio szybkie uzyskanie zbioru kombinacji. Tym, co podlega procesowi kombinowania mogą być słowa, jak to ma miejsce przy tworzeniu tekstu.

Jest też dla zrozumienia umysłu konieczne, by dostrzegać rolę kombinacji nie-werbalnych. Niech ją wyjaśni wypowiedź Einsteina z listu do Hadamarda, załączonego jako Dodatek II do książki o psychologii odkryć matematycznych.

«A. Wydaje mi się, że słowa albo język, czy to mówione czy pisane, nie odgrywają żadnej roli w moim procesie myślenia. Jednostki psychiczne służące mi, jak się zdaje, jako elementy myślenia są pewnymi znakami, niejasnymi obrazami, które można „swobodnie” odtwarzać i kombinować.

[...] Owa gra kombinacji wydaje się być istotną cechą twórczego myślenia — nie poprzedza jej żadna logiczna konstrukcja wyrażona bądź w słowach, bądź przy pomocy innego rodzaju znaków, które mogłyby być zakomunikowane innym.

(B) Wspomniane wyżej elementy są, w moim przypadku, typu wzrokowego, a niektóre — mięśniowego. Konwencjonalnych słów albo innych znaków poszukuję mozolnie dopiero w drugim stadium, kiedy wspomniana gra kombinacji jest dostatecznie rozbudowana. [...]

(E) Wydaje mi się, że to, co Pan nazywa pełną świadomością, jest przypadkiem granicznym, który nigdy w pełni nie ma miejsca.

Może ten opis być bardzo odległy od tego, czego doświadczają ludzie nie będący einsteinami, silnie uzależnieni od słów, za pomocą których się wyraża

obiegowe stereotypy. W każdym razie, w myśleniu ma się do czynienia z kombinacjami, czy to słów czy innych elementów, które kandydują do tego, żeby być rozwiązaniem, lub jakimś składnikiem rozwiązania, rozważanego problemu.

Dokonywanie wyboru wśród tych kandydatów wymaga ich testowania logicznego, które z reguły ma charakter niewerbalny. Testowanie polega na wyprowadzeniu z danej kombinacji takiego wniosku, który jest odpowiedzią na dany problem. Jeśli problem jest niebanalny, a droga do rozwiązania długa (gdyby ją rozpisać na kroki dowodowe na papierze), zaś kombinacja będąca przesłanką jest wysoce pomysłowa, a więc nieoczekiwana, a więc o bogatej zawartości informacyjnej, powstaje wtedy struktura logiczna, której się doświadcza jako piękna. Takie przeżycie piękna jest jakby skrótem czy syntetycznym ujęciem bogatej i zwartej logicznie struktury.

Istnieje też i odwrotność tego przeżycia, którą mimochodem wspomina Penrose, mianowicie doświadczenie brzydoty. Pojawia się ono w obliczu sprzeczności, do której doprowadza dana kombinacja idei lub w obliczu niemożności ujęcia związków logicznych (co zwiemy tradycyjnie *non sequitur*). I tutaj, podobnie jak w przeżyciu pozytywnym, na zasadzie jakby niewidzialnej błyskawicy, działa intuicja logiczna dotycząca konsekwencji, czasem bardzo odległych. Musi być ona procesem obliczeniowym wykonywanym przez jakiś procesor mózgowy, którego sprawne funkcjonowanie nie zależy od świadomości (a nawet może byłoby zakłócone przez świadomość, gdyby się ona włączyła). Tym natomiast, co jest sygnałem tego procesu wyłaniającym się na obszarze świadomości jest przeżycie piękna lub brzydoty, w zależności od tego, czy mamy bogactwo treści wraz ze spójnością logiczną, czy też ubóstwo lub niespójność.

Powyższa hipoteza wyjaśnia, dlaczego tak pojęta el-wrażliwość bywa zawodna jako kryterium (czy raczej źródło kryteriów) poprawności rozwiązania. Nie jest ona niemylna, gdyż nie jest niezawodny ów pozaświadomy proces obliczeniowy, który wylicza bogactwo informacyjne teorii i poprawność wiodącego do niej wnioskowania. Może się więc zdarzyć, że umysł się delektuje pięknem teorii, która w wyniku obliczeń logicznych wyłania się z pewnego zbioru kombinacji, ale jeśli w obliczeniach zdarzy się błąd (a zdolność do błędów jest wbudowana w procesory organiczne), to delektacja będzie przedwczesna i złudna. Podobnie może nas mylić, z powodu błędu obliczeniowego, poczucie brzydoty.

Jest zapewne i drugi powód omylności, ale ten będą brać pod uwagę tylko ci, którzy (jak obecny autor) wierzą w obiektywność piękna miast podpisywać się pod poglądem, że jest to tylko sprawa gustu. Wtedy sądy o pięknie

mogą być prawdziwe lub fałszywe. Dlaczego zdarzają się fałszywe, to osobne zagadnienie, nie do podejmowania w tym kontekście. W każdym razie, może się zdarzyć ten błąd, że piękna idea czy teoria zostanie przez kogoś oceniona jako brzydka, lub odwrotnie. To byłby drugi powód, dla którego el-wrażliwość nie jest arbitrem nieomylnym.

Jeśli obecne próby wyjaśnienia są trafne, to rysuje się interesujące zagadnienie, jak kształcić el-wrażliwość u adeptów nauki. Z dziejów nauki i nauczania znamy pewne metody. Na przykład, w wyniku treningu stosowanego w scholastyce odbierało się jako coś pięknego szarmonizowanie myśli z głosami autorytetów, zwłaszcza autorytetu Objawienia. W kulturze racjonalistycznej 17-go wieku tym, co miało największy powab była matematyzacja teorii. Z postawy scholastycznej brało się zalecenie treningu w respektowaniu autorytetów, z postawy racjonalistycznej — zalecenie treningu matematycznego (zaczęło ono się pojawiać już w łonie późniejszej scholastyki). Innym kryteriom estetycznym odpowiadają jeszcze inne zasady rozwijania el-wrażliwości.

Postawione wyżej zagadnienie dotyczy metod rozwijania el-wrażliwości jako składnika inteligencji człowieka. Podobne pod pewnym względem, lecz będące jeszcze większym wyzwaniem, jest zagadnienie konstrukcji sztucznej inteligencji obdarzonej el-wrażliwością. Korzystając ze stwierdzeń poczynionych wyżej na temat jej przejawiania się u ludzi, naszkicuję na koniec tych rozważań problem jej rekonstrukcji w maszynach.

5. Czy da się skonstruować elektroniczną el-wrażliwość?

Są autorzy, dla których odpowiedź twierdząca jest z góry przesądzona, nie ma więc pytania *CZY*, jest tylko pytanie *JAK*. Ci drudzy będą uważać dociekania typu *CZY* za marnotrawienie czasu i energii. Nie da się ich jednak uniknąć, gdy kieruje się rozważania do grona szerszego niż wymienieni autorzy. Nim przedstawię nieco szerzej to stanowisko, którego pionierem był Alan Turing, trzeba wyjaśnić, dlaczego w postawieniu pytania mowa jest tylko o maszynach elektronicznych, a nie o maszynach w ogólności.

Gdy pojęcie maszyny tak się uogólni, że – wzorem samego Leibniza – także organizmy obdarzone świadomością będzie się traktować jako pewien gatunek w klasie maszyn, to staje się trywialnie oczywiste, że istnieją maszyny obdarzone el-wrażliwością. Pytanie, czy taką maszynę potrafi skonstruować człowiek pozostaje aktualne, lecz wtedy przenosi się z informatyki do biotechnologii. Wraz z tym powstają problemy, które dalece by przekraczały tematykę tego artykułu i kompetencje jego autora. Gdy natomiast ograniczy się problem do maszyn elektronicznych w ich obecnym stanie, opartym

na technologii krzemowej, to pozostajemy na gruncie informatyki oraz tego, co można nazwać filozofią informatyki; w tej ostatniej mieści się zagadnienie stosunku *software'u* do *hardware'u*.

Taką filozofię uprawiał sam Alan Turing.² Żeby nie powtarzać jego tekstów, szeroko już znanych, skorzystam z wypowiedzi, która jest w jego duchu, a pochodzi z pracy aktualnie będącej w przygotowywaniu, uwzględniającej najnowszy stan techniki. Jest to pogląd Stephena Hawkinga.

«Jak dotąd, komputery nie wykazują nawet śladów inteligencji. Jest to o tyle zrozumiałe, że nawet najsilniejsze komputery są zbudowane prościej niż mózg dżdżownicy — zwierzęcia nie grzeszącego przecież inteligencją.

Nie zapominajmy jednak, że komputery podlegają tzw. prawu Gordona Moore'a, jednego z założycieli Intela, [...] że szybkość i stopień złożoności komputerów podwajają się co 18 miesięcy. Rzecz jasna, nie będzie to trwało wiecznie. Zmieni się w momencie, gdy układy elektroniczne osiągną stopień komplikacji porównywalny z ludzkim mózgiem.

Niektórzy uważają, że komputery nigdy się nie staną naprawdę inteligentne, cokolwiek by to miało oznaczać. Jestem innego zdania. Skoro skomplikowane cząsteczki chemiczne wchodzące w skład naszego mózgu uczyniły z nas istoty myślące, to tak samo efektem rosnącej komplikacji układów elektronicznych może być wykształcenie się inteligencji u komputerów.» Stephen Hawking, fragmenty pracy przygotowanej dla Rady Przełomu Tysiącleci przy Białym Domu; tekst za *New Perspectives Quarterly* w tłumaczeniu A. Włodarskiego zamieszczonym w *Gazecie Wyborczej*, 18-19 kwietnia, 1998, s. 15.

Przy konfrontacji obecnych rozważań z tekstem Hawkinga, cały problem ogniskuje się w pojęciu *wrażliwości*. Nie przypadkiem, wzorem Hada-marda i cytowanych przezeń autorów, wybrałem je na pojęcie kluczowe, opatrząc przymiotnikiem „estetyczno-logiczna” — dla wyodrębnienia tego jej gatunku, o którym mówili omawiani badacze.

Wrażliwość zakłada naturę organiczną, którą znamy jedynie w formie białkowej. Jeśli nie wykluczymy ewentualności, że życie organiczne i warunkowana przezeń *wrażliwość* mogłyby się oprzeć na związkach krzemu, to pojawi się w sferze możliwości i to, że procesory krzemowe posiadałyby *el-wrażliwość*. Nie wystarczyłoby jednak w tym celu, żeby je wyposażać

² Turing jest znakomitym przykładem łączenia maksymalnej kompetencji naukowej i technicznej z horyzontem filozoficznym. Tym dwom stronom jego twórczości odpowiadają dwie pionierskie prace na temat mechanicznej (czyli sztucznej) inteligencji. Jedna z nich to raport techniczny z roku 1948 pt. „Intelligent Machinery”, druga to artykuł zamieszczony w kwartalniku filozoficznym *Mind* (vol. 59, nr 236, October 1950). Obie znalazły się w *Collected Works of A. M. Turing*, w tomie *Mechanical Intelligence*.

w bardziej skomplikowane programy. Trzeba by uprzednio wywołać takie procesy chemiczne, które by uczyniły strukturę tego kryształu odpowiednio skomplikowaną. Do takiej rewolucji nie byłoby w stanie doprowadzić sami informatycy, musieliby oni współdziałać z biochemikami i biologami, którzy by dostarczyli odpowiedniego urządzenia (*hardware'u*).

Tej ewentualności nie bierze pod uwagę Hawking, który zdaje się zakładać (jak to czynił Turing), że strona materialna jest całkowicie obojętna, gdyż o wszystkim decyduje samo oprogramowanie. Gdyby tak było, to wspomniane interwencje biochemiczne nie byłyby potrzebne. Jeśli okażą się one konieczne, to stanowisko to będzie nie do utrzymania — chyba, że nada mu się mniej radykalną postać, która zasługuje na nazwę teorii *uniwersalnej symulacji*.³

Wedle tej teorii, świadomość, a więc i wrażliwość, jest jakby luksusem, który – z punktu widzenia efektywności rozwiązywania problemów – można w stosownych warunkach wyeliminować jako rzecz zbędną. Tak się złożyło w toku ewolucji, że rozwiązywanie problemów przez organizmy dokonuje się z udziałem świadomości i wrażliwości, ale dokładnie te same wyniki można będzie osiągnąć na drodze odpowiednich symulacji. Wtedy nie jest nie do pomyślenia, że do takich symulacji wystarczą elektroniczne urządzenia krzemowe, byleby ich wydajność sprostała wymaganiom skrajnie skomplikowanych programów symulujących.⁴

Te filozoficzne deliberacje nie są nieważne z technologicznego punktu widzenia. Pozwalają one uprzytomnić dylematy dotyczące strategii badawczych i kierunków rozwijania techniki. Inne powinny być wybory, a przynajmniej proporcje nakładów na badania i rozwój, jeśli racja jest po stronie teorii uniwersalnej symulacji, inne zaś o ile inteligencja, w tym tak ważny jej czynnik jak el-wrażliwość, wymaga bezwarunkowo podłoża organicznego. A jeśli

³ Tak się złożyło, że z dwóch „ojców komputerów”, jak się nazywa Alana Turinga i Johna von Neumanna, każdy reprezentował odmienne stanowisko w sprawie roli *hardware'u*. Podczas gdy dla Turinga była ona nieistotna, von Neumann wiązał z mózgiem możliwości inteligentnego działania biorące się z jego właściwości biologicznych. Dał temu wyraz m.in. *The Computer and the Brain*, Yale University Press, 1957. O tej różnicy traktuje Helmut Schnelle w studium „Turing Naturalized: Von Neumann's Unfinished Project” w: Rolf Herken (ed.), *The Universal Turing Machine. A Half-Century Survey*, Oxford University Press, 1988.

⁴ Jest to jedno ze stanowisk dyskutowanych przez Rogera Penrose (oznaczone u niego literą B) w książce *Shadows of the Mind. A Search for the Missing Science of Consciousness*, Oxford University Press, 1994.

to drugie, to, z kolei, trzeba wybierać między kontynuowaniem użytej przez ewolucję ścieżki białkowej a szukaniem ścieżek alternatywnych.

W dokonywaniu tego rodzaju wyborów nie wystarczą nauki szczegółowe, bo te dopiero czekają na program badawczy, który pozwoli uzyskać wyniki. Tworzenie takich programów jest wspólnym zadaniem dla przedstawicieli odpowiednich nauk oraz filozofów. Ci drudzy reprezentują sferę możliwości, ogarnianych przez wiedzę encyklopedyczną i rozum spekulatywny, pierwsi zaś sferę empirycznej rzeczywistości. Filozof, z reguły, nie umie doradzić, jaki wybór będzie właściwy, ale ukazując spektrum możliwości może pomóc w uniknięciu wyborów nie dość przemyślanych.

*

Co się tyczy el-wrażliwości, to rada filozofa co do badań nad jej wytworzeniem lub symulowaniem w maszynach byłaby następująca. Jak zostało zauważone wyżej w toku analizowania tej dyspozycji, jej istotnym składnikiem jest zdolność błyskawicznego i zarazem nieświadomego wyprowadzania konsekwencji. W tej dziedzinie istnieje rozległe i płodne pole symulacji komputerowych. Różnej mocy provery odpowiadałyby różnej długości „skokom” we wnioskowaniu. Rozwiązaniem, w którym kroki są możliwie najmniejsze jest formalizacja w stylu Hilberta, którą by symulował najsłabszy prover. W serii proverów coraz mocniejszych dochodziłoby się do takich, które by adekwatnie symulowały rozumowania najbardziej intuicyjne, a więc te idące najbardziej „na skróty”.

Czy tą drogą dojdziemy do wytworzenia lub symulowania owego istotnego składnika el-wrażliwości? Gdy brakuje mapy terenu, to żeby zbadać dokąd droga prowadzi, trzeba nią po prostu pójść. Niezależnie od tego, dokąd się dojdzie, będzie to niezwodnie wkład w zrobienie mapy.