Skłonnościowa interpretacja prawdopodobieństwa

Ostatnio w serii ROZPRAWY OBI ukazały się:

1. Maria Piesko, *Naukowa metafizyka Zygmunta Zawirskiego*, Tarnów 2004.
2. Anna Brożek, *Symetria w muzyce czyli o pierwiastku racjonalnym w komponowaniu dzieł muzycznych*, Tarnów 2004.
3. Paweł Polak, *Dynamika nauki. Filozoficzne aspekty modelowania rozwoju nauki przy pomocy układów dynamicznych*, Tarnów 2004.
4. Jacek Rodzeń, *Czy sukcesy nauki są cudem? Studium filozoficzno–metodologiczne argumentacji z sukcesu nauki na rzecz realizmu naukowego*, Tarnów 2005.
5. Teresa Obolevitch, *Problematyczny konkordyzm. Wiara i wiedza w myśli Włodzimierza S. Sołowjowa i Siemiona L. Franka*, Tarnów 2006.
6. Jerzy Dadaczyński, *Bernard Bolzano i idea logicyzmu*, Tarnów 2007.
7. Robert Piechowicz, *Logika, topologia, język. Relacja bliskości znaczeń na poziomie leksyki*, Tarnów 2007.
8. Wojciech P. Grygiel, *Zobaczyć kota Schrödingera*, Tarnów 2008.

Wojciech Załuski

Skłonnościowa interpretacja prawdopodobieństwa

OBI

Biblos

ISBN ??-????-???-?

Skład w systemie LaTeX: *OBI*
Projekt okładki: *Artur Piątek*

OŚRODEK BADAŃ INTERDYSCYPLINARNYCH
PRZY WYDZIALE FILOZOFICZNYM
PAPIESKIEJ AKADEMII TEOLOGICZNEJ
W KRAKOWIE
http://www.obi.opoka.org/

Wydawnictwo Diecezji Tarnowskiej
Biblos
Plac Katedralny 6, 33–100 Tarnów
tel. (0–14) 621–27–77
fax (0–14) 622–40–40
e-mail: biblos@wsd.tarnow.pl
http://www.biblos.pl/

Spis treści

Wykaz skrótów

SIP – skłonnościowa interpretacja prawdopodobieństwa

SIP_{lr} – skłonnościowa interpretacja prawdopodobieństwa *long run*

SIP_{sc} – skłonnościowa interpretacja prawdopodobieństwa *single case*

SIP_{scm} – skłonnościowa interpretacja prawdopodobieństwa *single case* metafizyczna

SIP_{scnm} – skłonnościowa interpretacja prawdopodobieństwa *single case* niemetafizyczna

Podziękowania

Niniejsza książka jest rozprawą doktorską, którą obroniłem na Wydziale Filozoficznym Papieskiej Akademii Teologicznej w Krakowie w czerwcu 2007 roku.

Chciałbym na początku podziękować wszystkim osobom, których życzliwa lektura tej rozprawy sprawiła, że odważyłem się ją opublikować: ks. prof. dr. hab. Stanisławowi Wszołkowi – promotorowi, ks. prof. dr. hab. Michałowi Hellerowi i dr. hab. Mieszkowi Tałasiewiczowi – recenzentom, a także dr. hab. Krzysztofowi Maślance. Wszystkie wyżej wymienione osoby dostarczyły mi wielu cennych uwag i komentarzy, które starałem się uwzględnić przygotowując rozprawę do druku.

W sposób szczególny chciałbym podziękować ks. prof. dr. hab. Stanisławowi Wszołkowi za okazaną mi ogromną życzliwość.

Pragnę także podziękować prof. dr. hab. Adamowi Groblerowi, który był promotorem mojej pracy magisterskiej „Teoria skłonności K. R. Poppera” w Instytucie Filozofii UJ w 2004 roku, stanowiącej pierwszy etap moich zmagań z teorią skłonnościową, a potem zachęcał mnie, aby owe zmagania kontynuować.

Dziękuję również Bartkowi Brożkowi za inspirujące rozmowy ‘wokół Poppera’, a także prof. dr. hab. Jerzemu Stelmachowi oraz prof. dr hab. Tomaszowi Gizbertowi-Studnickiemu za stworzenie mi w Katedrze Teorii i Filozofii Prawa UJ znakomitych warunków do pracy naukowej.

Pragnę także podziękować Gosi Szczerbińskiej-Polak i Pawłowi Polakowi za pomoc w przygotowaniu niniejszej książki do druku.

Książka ta nie powstałaby bez wsparcia osób mi najbliższych (w szczególności moich Rodziców), którym pragnę podziękować najserdeczniej.

Rzecz jasna, za wszelkie niedostatki niniejszej książki ponoszę wyłączną odpowiedzialność.

Wojciech Załuski

Wprowadzenie

Trudno dziś wyobrazić sobie możliwość uprawiania nauk przyrodniczych bez wykorzystania metod probabilistycznych. Metody te stosuje się nie tylko przy budowie wielu teorii z zakresu tych nauk, lecz niektóre spośród tych teorii, np. mechanika statystyczna i – w jeszcze głębszym sensie – mechanika kwantowa, są teoriami *stricte* probabilistycznymi. Metody probabilistyczne odgrywają ogromną rolę także w naukach społecznych: uprawiający ekonomię czy socjologię bardzo obficie korzystają z teorii racjonalnego wyboru i statystyki, a więc teorii, w których pojęcie prawdopodobieństwa zajmuje miejsce centralne. Rola metod probabilistycznych jest nie do przecenienia także w różnych gałęziach filozofii, zwłaszcza w filozofii nauki, gdzie okazują się one niezbędne do prowadzenia zaawansowanej analizy takich zagadnień, jak np. przyczynowość i konfirmacja teorii naukowych, oraz w filozofii umysłu, gdzie wykorzystuje się je np. w budowie modeli umysłu czy w analizie procesów uczenia się. Za pośrednictwem teorii racjonalnego wyboru pojęcie prawdopodobieństwa przenika także do etyki, filozofii społecznej, filozofii prawa i filozofii religii. Ta wszechobecność metod probabilistycznych inspiruje filozofów do jeszcze intensywniejszej niż dawniej refleksji nad samym pojęciem prawdopodobieństwa[1]. Refleksja ta zmierza do sformułowania odpowiedzi na py-

[1] Oczywiście, ekspansja metod probabilistycznych w nauce rozpoczęła się już w XIX wieku (m.in. W. Buniakowski zastosował te metody w statystyce i demografii, A. Quetelet w demografii i socjologii, F. Galton w psychologii

tanie o jego filozoficzny sens, czyli przede wszystkim na pytanie o to, czy prawdopodobieństwo jest miarą jakiejś własności świata czy też raczej naszych przekonań o świecie. W niniejszej pracy poddaję analizie jedną z kilku tego rodzaju odpowiedzi zaproponowanych przez filozofów. Chodzi o skłonnościową interpretację prawdopodobieństwa (w dalszej części używam na jej oznaczenie skrótu SIP). Jako pierwszy w sposób wyraźny interpretację tę sformułował Karl R. Popper. Następnie rozwijało ją wielu innych filozofów tak, że obecnie SIP nie jest, ściśle rzecz biorąc, jedną interpretacją prawdopodobieństwa, lecz raczej zbiorem interpretacji, które posiadają dwie cechy wspólne: są obiektywne, tj. zakładają, że prawdopodobieństwo jest miarą własności świata, a nie naszych przekonań o świecie; oraz opierają się na idei skłonności, tj. zakładają, że ową istniejącą niezależnie od naszych przekonań własnością świata, do której odnosi się prawdopodobieństwo, jest skłonność, tendencja, czy dyspozycja jakiegoś układu warunków do realizacji jakiegoś typu zdarzeń. Chcę jednak zaznaczyć już w niniejszym wprowadzeniu, że wkład filozofów w rozwój Popperowskiego ujęcia SIP jest – moim zdaniem – relatywnie niewielki. Popper nie tylko sformułował ideę prawdopodobieństwa jako skłonności w sposób klarowny, ale także – wyraźniej niż ktokolwiek inny – pokazał, jakie pożytki płyną z takiego ujęcia prawdopodobieństwa: argumentował, że SIP może stanowić klucz do zrozumienia mikroświata oraz – w konsekwencji – stać się jądrem metafizycznego programu badawczego organizującego pracę naukowców. Inni filozofowie jedynie doprecyzowali pewne sformułowania Poppera, podali różne układy aksjomatów dla skłonności, oraz podjęli próbę wykorzystania SIP w różnych gałęziach nauki. Tytuł niniejszej pracy powinien więc właściwie brzmieć „Skłonnościowa

---

i antropologii, K. Pearson w biologii, J. Maxwell, L. Boltzmann i J. Gibbs w fizyce statystycznej), ale wszechobecne stały się one dopiero w wieku XX; por. na ten temat np. Gigerenzer et al. 1989, Hacking 2005b, czy Heller 2006d, s. 68-70.

interpretacja prawdopodobieństwa Karla R. Poppera”. Ponieważ jednak tytuł ten mógłby chyba w jeszcze większym stopniu niż obecny sugerować, że istnieją jakieś inne wersje SIP zasadniczo odrębne od Popperowskiej, ostatecznie zrezygnowałem z niego. W każdym razie, chcę podkreślić, że SIP jest przede wszystkim interpretacją Popperowską.

Plan pracy jest następujący. W rozdziale 1. omawiam najważniejsze interpretacje rachunku prawdopodobieństwa. Opis tych interpretacji ma stanowić tło pozwalające wyraźniej uchwycić specyfikę SIP. Wskazuję kilka kryteriów oceny interpretacji prawdopodobieństwa i próbuję dokonać wstępnej oceny wszystkich interpretacji za wyjątkiem SIP. Ocenę tej interpretacji formułuję w Zakończeniu. W rozdziale 2. analizuję trzy podstawowe argumenty za przyjęciem tej interpretacji: pewien eksperyment myślowy Poppera mający pokazać, że SIP jest interpretacją bardziej przekonującą niż interpretacja częstościowa; argument, iż SIP najlepiej z wszystkich interpretacji „wpisuje się” w indeterministyczną wizję świata; oraz argument, iż SIP pozwala rozwiązać niektóre konkretne problemy filozoficzne i naukowe (przede wszystkim problem przypisywania prawdopodobieństw zdarzeniom jednostkowym, problem interpretacji mechaniki kwantowej oraz problem stabilności względnych częstości). W tym rozdziale poddaję także krytyce tezę, iż SIP może pomóc w rozwiązaniu problemu interpretacji pojęcia dostosowania, który pojawia się na gruncie teorii ewolucji. W rozdziale 3. przedstawiam główne zarzuty przeciw SIP: zarzut metafizyczności, zarzut wtórności, oraz zarzut, iż na gruncie tej interpretacji nie można nadać jasnego sensu tzw. prawdopodobieństwom „odwrotnym”. Jak już wspomniałem, w Zakończeniu próbuję dokonać oceny SIP, „ważąc” argumenty za jej przyjęciem i argumenty za jej odrzuceniem analizowane w rozdziałach 2. i 3.

Rozdział 1
Interpretacja skłonnościowa na tle innych interpretacji prawdopodobieństwa

1.1. Matematyczna eksplikacja pojęcia prawdopodobieństwa

Generalnie rzecz biorąc, wśród osób badających różne aspekty pojęcia prawdopodobieństwa panuje zgoda w kwestii jego aspektów matematycznych. Większość z nich przyjmuje tzw. standardowy – zaproponowany przez Andrieja N. Kołmogorowa w 1933 roku w dziele *Grundbegriffe der Wahrscheinlichkeitsrechnung* – układ aksjomatów, który definiuje prawdopodobieństwo. Tę aksjomatyczną definicję prawdopodobieństwa można traktować jako matematyczną eksplikację przed-teoretycznego rozumienia prawdopodobieństwa, zgodnie z którym prawdopodobieństwo jest miarą możliwości zachodzenia zdarzeń i/lub miarą niepewności co do tego, czy dane zdarzenie zajdzie czy nie. Przejdę teraz do omówienia tej definicji[1].

Z czysto matematycznego punktu widzenia prawdopodobieństwo jest nieujemną, unormowaną, przeliczalnie addytywną miarą określoną na pewnym rodzaju algebry Boole'a. W kontekście definiowania prawdopodobieństwa do interpretacji algebry Boole'a używa się zazwyczaj języka teorii mnogości. Symbolem Ω oznacza się zbiór pełny. Zbiór ten – nazywany „przestrzenią zdarzeń elementarnych" – zawiera wszystkie możliwe wyniki doświadczenia losowego, czyli tzw. zdarzenia elementarne (zdarze-

[1] Omówienie to oparte jest przede wszystkim na: Jakubowski, Sztencel 2004, s. 9-11.

nia te muszą być wyodrębnione w taki sposób, aby wykluczały się wzajemnie, i aby w każdym doświadczeniu losowym zachodziło dokładnie jedno z nich). Prawdopodobieństwo nie odnosi się jednak do zdarzeń elementarnych, lecz do zdarzeń rozumianych jako podzbiory przestrzeni zdarzeń elementarnych Ω (podzbiory te tworzą pewną rodzinę F, której dokładniejszą charakterystykę podaję poniżej). Każde z tych zdarzeń jest utożsamiane z podzbiorem złożonym z tych zdarzeń elementarnych, które mu „sprzyjają" (np. jeśli wynikiem doświadczenia jest zdarzenie elementarne $\omega \in A$, wtedy powiemy, że zaszło zdarzenie A). Wśród tych zdarzeń dwa są wyróżnione: wspomniany już zbiór pełny Ω – zdarzenie zachodzące zawsze, które nazywa się „zdarzeniem pewnym", oraz zbiór pusty $\emptyset$ – zdarzenie nie zachodzące nigdy, które nazywa się „zdarzeniem niemożliwym". Należy podkreślić fakt, że na podzbiorach Ω tworzących rodzinę zdarzeń F funkcję prawdopodobieństwa można opisać tylko wtedy, gdy rodzina ta jest σ-ciałem (σ-addytywną algebrą Boole'a) podzbiorów zbioru Ω, tzn. tylko wtedy, gdy spełnia następujące warunki:

(*W1*) $\Omega \in F$;

(*W2*) Jeśli $A \in F$, to $A' \in F$;

(*W3*) Jeśli $A_i \in F$ dla $i = 1, 2, \ldots$, to

$$\bigcup_{i=1}^{\infty} A_i \in F.$$

Jeśli warunek *W3* zastąpi się słabszym:

(*W3'*) Jeśli $A, B \in F$, to $A \cup B \in F$,

to warunki *W1-W3'* będą definiować ciało podzbiorów zbioru Ω. Jak widać, warunki domknięcia dla σ-ciała zbiorów są silniejsze niż dla ciała zbiorów: rodzina zbiorów F jest σ-ciałem zbiorów, jeśli jest domknięta na operacje brania dopełnień i przeliczalnych sum, a nie tylko na operacje brania dopełnień i skończonych sum. Ten mocniejszy warunek addytywności, tj. warunek *W3*, jest konieczny, jeśli chce się wprowadzić aksjomat przeliczalnej addytywności (o którym jest mowa poniżej).

Na σ-ciele podzbiorów zbioru Ω określa się odpowiednią funkcję – funkcję prawdopodobieństwa. *Prawdopodobieństwo jest dowolną funkcją rzeczywistą P, której zbiór argumentów tworzy σ-ciało F podzbiorów zbioru Ω, i która spełnia następujący układ aksjomatów A1-A3*:

($A1$) *Nieujemność*. $P(A) \geqslant 0$, dla każdego $A \in F$;

($A2$) *Unormowanie*. $P(\Omega) = 1$;

($A3$) *Przeliczalna addytywność*. Jeśli $A_i \in F$, $i = 1, 2, \ldots$ oraz $A_i \cap A_j = \emptyset$ dla $i \neq j$, to

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Podsumowując, matematyczny model doświadczenia losowego to struktura (Ω, F, P), zwana „przestrzenią probabilistyczną”, gdzie P jest nieujemną, unormowaną, przeliczalnie addytywną miarą określoną na σ-ciele F podzbiorów zbioru Ω, tj. zbioru zdarzeń elementarnych.

Godne podkreślenia jest to, że z punktu widzenia matematycznego prawdopodobieństwo jest funkcją, a nie wartością funkcji. Znaczy to, że wyrażenie „$P(A) = r$” powinno się czytać „wartość funkcji prawdopodobieństwa P dla argumentu A wynosi r”. W dalszej części pracy będę jednak często posługiwał się skrótowymi sformułowaniami w rodzaju „prawdopodobieństwo A wynosi r”.

Na podstawie powyższych aksjomatów można udowodnić m.in. następujące własności prawdopodobieństwa *Wł1-Wł7*. Jeśli (Ω, F, P) jest przestrzenią probabilistyczną i A, B, $A_1, A_2, \ldots, A_n \in F$, to:

(*Wł1*) *Prawdopodobieństwo zdarzenia niemożliwego jest równe zero*. $P(\emptyset) = 0$.

($W\!ł2$) *Skończona addytywność*[2]. Jeśli $A_1, A_2, \ldots, A_n$ wykluczają się wzajemnie, czyli $A_i \cap A_j = \emptyset$ dla $i \neq j$, $1 \leqslant i, j \leqslant n$, to

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

($W\!ł3$) *Prawdopodobieństwo zdarzenia przeciwnego.* $P(A') = 1 - P(A)$.

($W\!ł4$) Jeśli $A \subset B$, to $P(B \backslash A) = P(B) - P(A)$.

($W\!ł5$) *Monotoniczność*. Jeśli $A \subset B$, to $P(A) \leqslant P(B)$.

(*Wł6) Prawdopodobieństwo żadnego zdarzenia nie przekracza jedności.* $P(A) \leqslant 1$.

($W\!ł7$) *Wzór na prawdopodobieństwo sumy zdarzeń.* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Aksjomaty (*A1*)-(*A3*) dotyczą prawdopodobieństwa absolutnego, które – jak widzieliśmy – jest funkcją jednoargumentową. Za jego pomocą można zdefiniować prawdopodobieństwo warunkowe $P(A|B)$, które jest funkcją dwuargumentową. Prawdopodobieństwem zajścia zdarzenia A pod warunkiem zajścia zdarzenia B nazywamy funkcję $P(A/B) = P(A \cap B)/P(B)$ (dla $P(B) > 0$). Niekiedy jednak prawdopodobieństwo warunkowe traktuje się jako pojęcie podstawowe rachunku prawdopodobieństwa, tzn. pojęcie, które definiuje się aksjomatycznie, a nie wprowadza – jak uczynił to Kołmogorow – za pomocą definicji[3]. Jednym z przekonujących argumentów za tym, aby za podstawowe uważać właśnie pojęcie prawdopodobieństwa warunkowego, a nie

[2] Rachunek prawdopodobieństwa z aksjomatem skończonej addytywności zamiast przeliczalnej addytywności Kołmogorow nazywał „elementarną teorią prawdopodobieństwa". W koniunkcji z tzw. aksjomatem ciągłości aksjomat skończonej addytywności jest równoważny aksjomatowi przeliczalnej addytywności. Aksjomat przeliczalnej addytywności występuje w dziele Kołmogorowa Grundbegriffe der Wahrscheinlichkeitsrechnung właśnie w takiej postaci – koniunkcji aksjomatu skończonej addytywności i aksjomatu ciągłości; por. Kolmogorow 1956.

[3] Tak ujmują prawdopodobieństwo warunkowe np. Popper (zob. np. Popper 2002b, s. 329-364) i W. Spohn (zob. np. Spohn 1986).

absolutnego, jest to, że, jak pokażę w dalszej części pracy, każda z prezentowanych interpretacji prawdopodobieństwa relatywizuje wypowiedzi probabilistyczne: klasyczna – do zbioru możliwości, logiczna – do zbioru świadectw, subiektywna – do wiedzy podmiotu, częstościowa – do klasy odniesienia, zaś skłonnościowa – do układu eksperymentalnego (za wyjątek można uznać jedynie pewien wariant tej ostatniej zakładający, że wartości prawdopodobieństw odnoszą się do stanów świata). Konsekwencją takiego ujęcia jest m.in. to, że prawdopodobieństwa warunkowe są zdefiniowane nawet wtedy, gdy zdarzenie warunkujące ma prawdopodobieństwo równe zero[4]. Ujęcie to zakłada więc, że prawdopodobieństwa są zawsze relatywne, tj. że żadnemu zdarzeniu nie można przypisać prawdopodobieństwa, nie wskazując *względem czego* ma ono daną wartość prawdopodobieństwa. Innymi słowy, ujęcie to zakłada, że prawdopodobieństwo jest pierwotnie funkcją dwuargumentową, a nie jednoargumentową. Innym pojęciem wprowadzonym do rachunku prawdopodobieństwa za pomocą definicji, a nie dodatkowego aksjomatu, jest niezależność probabilistyczna. O dwóch zdarzeniach A i B mówi się, że są niezależne probabilistycznie wtedy i tylko wtedy, gdy $P(A/B) = P(A)$ (dla $P(B) > 0$) (z uwagi na definicję prawdopodobieństwa warunkowego warunek ten jest równoważny warunkowi $P(A \cap B) = P(A)P(B)$; to równoważne ujęcie pokazuje jasno, że niezależność probabilistyczna jest relacją symetryczną). Wiele ważnych twierdzeń rachunku prawdopodobieństwa zakłada niezależność probabilistyczną zdarzeń. Wśród nich szczególnie istotne z filozoficznego punktu widzenia

[4] Rozważmy następujący przykład. Wybieramy losowo liczbę rzeczywistą z przedziału [0, 1]. Prawdopodobieństwo tego, że wybierzemy albo 1/4, albo 3/4 jest równe zero. Nie mamy jednak wątpliwości, że prawdopodobieństwo tego, że wybierzemy 1/4 pod warunkiem, że wybraliśmy albo 1/4, albo 3/4, wynosi 1/2. W ujęciu tradycyjnym wartość tego prawdopodobieństwa jest jednak niezdefiniowana. Por. Hájek 2003b (w całości poświęcony krytyce tradycyjnego ujęcia prawdopodobieństw warunkowych) oraz Hájek 2006, s. 33-41.

są różne warianty prawa wielkich liczb, które, mówiąc swobodnie, zakłada, że w miarę powtarzania doświadczenia losowego, względna częstość występowania wyniku A generowanego w tym doświadczeniu będzie zbliżać się do prawdopodobieństwa tego wyniku[5].

W powyższym omówieniu prawdopodobieństwo zostało ujęte jako funkcja określona na spełniającym określone warunki zbiorze zdarzeń F, tj. algebra Boole'a została zinterpretowana w rachunku zbiorów. Nic nie stoi jednak na przeszkodzie, aby prawdopodobieństwo ująć jako funkcję określoną nie na zbiorze zdarzeń, lecz na spełniającym określone warunki zbiorze zdań Z. Wystarczy dane zdarzenie zastąpić jego opisem, a operacje teoriomnogościowe operacjami logicznymi. Zbiór taki musi być domknięty na operacje klasycznego rachunku zdań. W ujęciu, w którym algebra Boole'a jest interpretowana w rachunku zdań, otrzymujemy następujący układ aksjomatów dla prawdopodobieństwa (ponieważ przyjąłem założenie, że zbiór zdań Z jest skończony, chodzi o układ ze skończoną addytywnością):

(*A1*) *Nieujemność*. $P(A) \geqslant 0$, dla każdego $A \in Z$;

(*A2*) *Unormowanie*. Jeśli T jest prawdą logiczną, to (w logice klasycznej) $P(T) = 1$.

(*A3*) *Skończona addytywność*. Jeśli zdania $A_1, A_2, \ldots, A_n \in Z$ wykluczają się wzajemnie, tj. jeśli żadne dwa zdania $A_i, A_j \in Z$, dla $i \neq j$, nie są równocześnie prawdziwe, to

$$P(\bigvee_{i=1}^{n} A_i) = \sum_{n=1}^{n} P(A_i).$$

---

[5] Jeden z najprostszych wariantów prawa wielkich liczb, tzw. klasyczne prawo wielkich liczb (twierdzenie Bernoullego), stwierdza, że jeśli w n powtórzeniach określonych warunków notujemy m wystąpień zdarzenia, którego prawdopodobieństwo jest równe p, to dla dowolnego $\epsilon > 0$, $P\{|m/n - p| < \epsilon\} \rightarrow 1$, gdy n wzrasta nieograniczenie; por. Feller 1977, rozdział VI, 4.

Wydaje się, że interpretując prawdopodobieństwa obiektywnie (tzn. jako miarę własności świata) bardziej naturalnie jest mówić o prawdopodobieństwie zdarzeń, zaś interpretując prawdopodobieństwa epistemicznie (tzn. jako miarę własności naszych przekonań) – o prawdopodobieństwie zdań[6].

Wspomniałem na wstępie niniejszej sekcji, że, generalnie rzecz biorąc, wśród osób zajmujących się pojęciem prawdopodobieństwa panuje zgoda w kwestii jego aspektów matematycznych. Klauzulę „generalnie rzecz biorąc" uzasadnia to, że wprawdzie większość z nich akceptuje układ aksjomatów Kołmogorowa, niemniej niektórzy przyjmują układy odmienne. Różnice między tymi układami a układem Kołmogorowa mogą polegać m.in. na porzuceniu aksjomatu przeliczalnej addytywności, aksjomatu skończonej addytywności[7], aksjomatu unormowania, aksjomatu nieujemności, na dopuszczeniu jako wartości prawdopodobieństwa liczb definiowanych w ramach analizy niestandardowej (tzn. dodatnich, ale mniejszych niż każda dodatnia liczba rzeczywista, czyli nieskończenie małych), czy na dopuszczeniu rozmytych (*vague*) lub nieokreślonych (*indeterminate*) prawdopodobieństw[8]. Ponadto, funkcja prawdopodobieństwa odnosząca się do mikroświata różni się od funkcji prawdopodobieństwa zdefiniowanej w ramach aksjomatyki Kołmogorowa m.in. tym, że: (i) algebra przestrzeni zdarzeń, na której jest określona, nie jest algebrą Boole'a, lecz niedystrybutywną algebrą kraty; (ii) nie obowiązuje dla niej generalnie aksjomat addytywności, tj. $P(A \cup B) \neq P(A) + P(B) - P(A \cap B)$ (co wiąże się z tym, że dla tzw. obserwabli kanonicznie sprzężonych, a więc nie komutują-

---

6 Pojęcia obiektywnej i epistemicznej interpretacji prawdopodobieństwa omawiam dokładnie w sekcji 1.8 tego rozdziału.

7 Taką teorią jest np. teoria Dempstera-Shafera formalizująca pojęcie „wagi świadectw" (teoria ta jest szczególnie użyteczna w rozważaniach prawniczych), czy teoria „prawdopodobieństw Bacona", zgodnie z którą prawdopodobieństwo koniunkcji jest równe prawdopodobieństwu tego z jej argumentów, dla którego jest ono najmniejsze.

8 Por. Hájek 2001, s. 372-376.

cych ze sobą, i tym samym – zgodnie z zasadą nieoznaczoności – nie dających się zmierzyć równocześnie z dowolną precyzją, mechanika kwantowa nie pozwala określić tzw. łącznych prawdopodobieństw (*joint probabilities*))[9]. Należy także wspomnieć, że podejmowane są próby matematycznego ujęcia mechaniki kwantowej za pomocą tzw. swobodnej teorii prawdopodobieństwa (*free probability theory*) Dana Voiculescu, która tym m.in. różni się od teorii standardowej, że w jej ramach pojęcie niezależności zdarzeń zostaje zastąpione pojęciem swobody (ma to związek z faktem, że w algebrze nieprzemiennej, w których swobodna teoria prawdopodobieństwa ma zastosowanie, nie można mówić o zdarzeniach w zwykłym sensie, tj. jako zlokalizowanych w konkretnym miejscu i czasie)[10].

Fakt istnienia odmiennych aksjomatyzacji prawdopodobieństwa, których kryteria oceny są w dodatku dość nieostre (o czym będzie mowa w sekcji 1.2.2), komplikuje zadanie wyboru interpretacji prawdopodobieństwa: interpretacja, która jest adekwatna na gruncie jednej aksjomatyki, niekoniecznie musi adekwatna na gruncie innej. Pamiętając o owych alternatywnych ujęciach w stosunku do ujęcia Kołmogorowa, należy jednak raz jeszcze podkreślić, że zarówno w naukach przyrodniczych jak i społecznych korzysta się przede wszystkim z aksjomatyki Kołmogorowa. Matematyczną eksplikacją prawdopodobieństwa zaproponowaną przez Kołmogorowa można więc uznać za podstawową.

1.2. Pojęcie interpretacji prawdopodobieństwa

1.2.1. Czym jest interpretacja prawdopodobieństwa?

Z rozważań prowadzonych w poprzedniej sekcji wynika, że w kontekście analizy prawdopodobieństwa zadanie matematyka

[9] Por. np. Baggott 2004, s. 76-77, Dickson 1998, s. 1-10, Redhead 2002, s. 1-42, Sklar 1992, s. 196, Wilce 2006.

[10] Por. Biane 1998, Heller 2006a, Nica, Speicher, Voiculescu 2004.

sprowadza się do zbadania własności określonej struktury matematycznej – przestrzeni probabilistycznej. Częścią tego zadania jest wybór zbioru aksjomatów prawdopodobieństwa oraz wyprowadzenie z tych aksjomatów za pomocą reguł wnioskowania twierdzeń rachunku prawdopodobieństwa. Te aksjomaty i twierdzenia definiują aksjomatycznie symbol prawdopodobieństwa „*P*", tzn. „rządzą" posługiwaniem się tym symbolem. Matematyczne ujęcie prawdopodobieństwa, które nazwałem „matematyczną eksplikacją przed-teoretycznego pojęcia prawdopodobieństwa", traktujące prawdopodobieństwo jako niezdefiniowane *explicite* pojęcie pierwotne, nie pozwala jednak odpowiedzieć na pytanie o filozoficzny sens tego pojęcia, czyli przede wszystkim na pytanie o to, czy prawdopodobieństwo jest miarą jakiejś własności świata, która jest niezależna od naszej świadomości, czy też raczej jest miarą naszych przekonań[11]. Precyzyjną odpowiedź na to pytanie można nazwać „filozoficzną eksplikacją przed-teoretycznego pojęcia prawdopodobieństwa" lub „interpretacją prawdopodobieństwa". W dalszej części pracy będę posługiwał się tym drugim terminem. W tym miejscu chciałbym jeszcze odnieść się do pewnego zastrzeżenia, jakie przeciw używaniu tego terminu zgłosił Alan Hájek. Otóż Hájek twierdzi, że nazwa „interpretacja prawdopodobieństwa" jest niefortunna, gdyż wszystkie wielkości, które bada się w ramach abstrakcyjnej teorii miary, a które nie mają nic wspólnego z prawdopodobieństwem (np. masa, długość, powierzchnia, objętość), spełniają (o ile są unormowane) aksjomaty Kołmogorowa, nikt ich jednak nie uważa za interpretacje prawdopodobieństwa, gdyż pełnią inną niż prawdopodobieństwo rolę w naszym aparacie po-

[11] Jak pisał R. von Mises, „system aksjomatów nie może zastąpić próby rozjaśnienia i rozgraniczenia pojęcia prawdopodobieństwa. Stanie się to oczywiste, gdy pomyślimy o przykładzie kostki do gry lub monet, gdzie [...] matematyczne trudności nie istnieją, lub raczej gdzie ich rozwiązanie jest bezpośrednie, bez korzystania z matematycznej teorii zbiorów (von Mises 1957, s. 99)".

jęciowym[12]. Wydaje mi się, że zastrzeżenie Hájeka jest trafne tylko wtedy, gdy zakłada się, że interpretacja prawdopodobieństwa jest interpretacją układu aksjomatów Kołmogorowa (czy ewentualnie jakiegoś innego układu aksjomatów prawdopodobieństwa) w takim sensie, jaki pojęciu interpretacji nadaje się w ramach teorii modeli (tj. gdy zakłada się, że interpretacja prawdopodobieństwa jest przypisaniem takich znaczeń terminom pierwotnym systemu formalnego, jakim jest standardowy rachunek prawdopodobieństwa, które przekształcają aksjomaty i twierdzenia tego systemu w zdania prawdziwe). Zgodnie jednak z zaproponowaną przeze mnie definicją, interpretacja prawdopodobieństwa jest *filozoficzną eksplikacją przed-teoretycznego pojęcia prawdopodobieństwa* (musi więc *ex definitione* pozostawać w pojęciowym związku z przed-teoretycznym rozumieniem prawdopodobieństwa jako miary możliwości zachodzenia zdarzeń i/lub miary niepewności co do tego, czy dane zdarzenie zajdzie czy nie), a zgodność z układem aksjomatów Kołmogorowa stanowi po prostu jedno z kilku kryteriów oceny jej akceptowalności. Relacja między moim rozumieniem interpretacji prawdopodobieństwa i rozumieniem wynikającym z teorii modeli jest więc następująca: filozoficzna eksplikacja przed-teoretycznego pojęcia prawdopodobieństwa (interpretacja prawdopodobieństwa w moim rozumieniu) nie jest akceptowalna, jeśli nie jest interpretacją w sensie teorii modeli (tj. jeśli nie spełnia układu aksjomatów Kołmogorowa).

Zanim przejdę do bliższego omówienia innych kryteriów oceny interpretacji prawdopodobieństwa, chciałbym jeszcze poruszyć kwestię istotności problemu interpretacji prawdopodobieństwa. Otóż można niekiedy spotkać się z opinią, że problem ten nie jest istotny. Wsparciem dla tej opinii ma być fakt, że większość naukowców wykonuje z powodzeniem niezbędne dla ich badań obliczenia probabilistyczne, nie usiłując wcale zgłębić tego, czym „w istocie” jest prawdopodobieństwo. Niekiedy

[12] Por. Hájek 2003a.

opinię tę formułuje się mocniej: powiada się, że pytanie o filozoficzny sens prawdopodobieństwa nie tylko nie jest istotne, ale że jest przykładem pseudo-problemu filozoficznego. Zarzut ten jednak należy uznać za powierzchowny. Po pierwsze, zadanie filozofii nauki polega m.in. właśnie na rozjaśnianiu pojęć, którymi posługują się naukowcy. Kwestionowanie sensowności tego rodzaju analiz oznacza *de facto* kwestionowanie samej sensowności uprawiania filozofii nauki. Takie analizy są niezbędne w kontekście większości „obciążonych filozoficznie” pojęć, którymi posługują się naukowcy, tym bardziej więc w kontekście pojęcia prawdopodobieństwa, które odgrywa we współczesnej nauce rolę zupełnie kluczową. Nie dysponując klarownym poglądem na temat tego, czym „w istocie” jest prawdopodobieństwo, trudno byłoby zrozumieć, dlaczego rachunek prawdopodobieństwa jest tak skuteczny w rozwiązywaniu problemów naukowych i filozoficznych. Po drugie, pogląd tego rodzaju ma nie tylko walor teoretyczny, ale i praktyczny: pozwala wyostrzyć rozumienie sensu wyników konkretnych badań, w których wykorzystuje się rachunek prawdopodobieństwa, oraz może wpłynąć na sposób, w jaki naukowcy będą przeprowadzać eksperymenty służące zebraniu najbardziej dla nich użytecznych informacji. Ujęcie aksjomatyczne nie może w tym kontekście zastąpić analizy pojęciowej. Po trzecie, należy pamiętać, że rachunek prawdopodobieństwa zawiera aksjomaty i twierdzenia pozwalające obliczać prawdopodobieństwa rozmaitych zdarzeń na podstawie znanego prawdopodobieństwa innych zdarzeń. Na podstawie aksjomatyki tego rachunku wiemy tylko, że dla zbioru pełnego, który interpretujemy jako zdarzenie pewne, prawdopodobieństwo przyjmuje wartość 1, zaś dla zbioru pustego, który interpretujemy jako zdarzenie niemożliwe, wartość 0. W związku z tym pojawia się pytanie, w jaki sposób powinno się określać wartości prawdopodobieństw zdarzeń innych niż zdarzenia pewne i niemożliwe? Otóż ich ustalenie musi się dokonywać na podstawie jakiegoś rozumienia tego, czym jest prawdopodobieństwo, a z pewnością jest pożądane, aby owo ustalenie dokonywało się z jasną świadomością przyjmowanych założeń filozoficznych. Wynika stąd, że interpretację prawdopodobieństwa można rozumieć także jako charakterystykę epistemologicznych podstaw prawdopodobieństwa (tj. jako określenie metody wyznaczania wyjściowych rozkładów prawdopodobieństwa), które są pomijane w czysto aksjomatycznym omówieniu tego pojęcia.

1.2.2. Kryteria oceny interpretacji prawdopodobieństwa

Przejdę teraz do omówienia kryteriów oceny interpretacji prawdopodobieństwa. Wesley C. Salmon wyróżnił trzy takie kryteria: dopuszczalności, stwierdzalności i stosowalności[13]:

(1) *Kryterium dopuszczalności* (*admissibility*). Interpretacja prawdopodobieństwa powinna spełniać aksjomaty i twierdzenia rachunku prawdopodobieństwa, tj. powinna być zgodna z matematyczną eksplikacją prawdopodobieństwa. Innymi słowy, obiekty, co do których przyjmuje się na gruncie danej interpretacji, że ich miarą jest prawdopodobieństwo, powinny być obiektami spełniającymi aksjomatyczną definicję prawdopodobieństwa lub powinny prowadzić w sposób wyraźny do skonstruowania obiektów, które spełniają tę definicję[14]. Kryterium to, o czym już wspominałem, implikuje, że interpretacja prawdopodobieństwa musi być interpretacją w sensie, jaki pojęciu interpretacji nadaje się w ramach teorii modeli.

(2) *Kryterium stwierdzalności* (*ascertainability*). Interpretacja prawdopodobieństwa powinna umożliwiać – przynajmniej co do zasady – określanie wartości prawdopodobieństw.

(3) *Kryterium stosowalności* (*applicability*). Interpretacja prawdopodobieństwa powinna umożliwiać określanie takich wartości prawdopodobieństwa, które mogą służyć jako „przewodnik

[13] Por. Salmon 1967, s. 63-65.

[14] Por. Suppes 1974b, s. 762-763.

w życiu"[15], tj. które mogą być podstawą racjonalnych decyzji. Kryterium to rozumiem jako stwierdzające, iż interpretacja prawdopodobieństwa powinna umożliwiać dokonywanie oszacowań wartości prawdopodobieństw przydatnych z punktu widzenia teorii racjonalnego wyboru[16].

Sądzę, że do trzech powyższych kryteriów wskazanych przez Salmona – kryteriów rozsądnych i nie wymagających chyba uzasadnienia – warto dodać jeszcze dwa, które będę nazywał „kryterium jasności" i „kryterium uniwersalności"[17].

(4) *Kryterium jasności*. Interpretacja prawdopodobieństwa powinna być sformułowana za pomocą dobrze zdefiniowanych pojęć[18].

(5) *Kryterium uniwersalności*. Jest pożądane, aby interpretacja prawdopodobieństwa była adekwatna dla możliwie wielu kontekstów, w jakich formułuje się w dyskursie potocznym i naukowym wypowiedzi probabilistyczne.

Wymienione kryteria wskazują na pożądane cechy, jakich oczekuje się od akceptowalnej interpretacji prawdopodobieństwa. Niestety nie stanowią one jakiegoś algorytmu pozwalającego automatycznie eliminować „nieakceptowalne" i zatwierdzać „akceptowalne" interpretacje prawdopodobieństwa. Jest tak co

---

[15] Ten wymóg nawiązuje do słynnej wypowiedzi biskupa J. Butlera (1692-1752), że „Probability is the very guide of life".

[16] Salmon pisze, że kryterium to nakazuje szukać takiego pojęcia prawdopodobieństwa, które „będzie mieć praktyczne, predykcyjne znaczenie (1967, s. 64)".

[17] Sformułowanego wyżej wymogu, zgodnie z którym interpretacja prawdopodobieństwa powinna zachowywać związek pojęciowy z przed-teoretycznym rozumieniem prawdopodobieństwa, nie traktuję jako kryterium oceny prawdopodobieństwa, lecz jako część samej definicji interpretacji prawdopodobieństwa.

[18] O kryterium jasności wspomina Hájek, kryterium uniwersalności jest moją propozycją. Jak słusznie zauważa Hájek, wymóg, aby interpretacja prawdopodobieństwa była jasna, nie jest – w odróżnieniu od pozostałych kryteriów – specyficzny dla interpretacji prawdopodobieństwa, lecz odnosi się do każdej teorii filozoficznej; por. Hájek 2003a.

najmniej z dwóch powodów. Po pierwsze, niektóre kryteria (2, 3, 4, 5) są sformułowane w sposób nieostry. Dana interpretacja może w mniejszym lub większym stopniu umożliwiać przypisywanie wartości prawdopodobieństw, może być mniej lub bardziej stosowalna, może być mniej lub bardziej precyzyjnie sformułowana, lub może być mniej lub bardziej uniwersalna. Po drugie, kryterium 1 jest wprawdzie ostre, ale za to zrelatywizowane do danego układu aksjomatów prawdopodobieństwa. Kłopot w tym, że nie istnieje jeden, powszechnie przyjmowany układ aksjomatów prawdopodobieństwa. Jak już wspominałem, wprawdzie układ Kołmogorowa jest uważany za standardowy, nie można go jednak nazwać „powszechnie przyjmowanym". Czy oznacza to, że spełnianie kryterium dopuszczalności nie jest warunkiem koniecznym uznania danej interpretacji prawdopodobieństwa za akceptowalną? Taka konkluzja byłaby, jak sądzę, nietrafna. Wydaje mi się, że kryterium dopuszczalności może pozostać takim warunkiem, o ile będzie się go ujmować nie-rygorystycznie, tj. nie jako wymóg spełniania wszystkich aksjomatów i twierdzeń standardowego rachunku prawdopodobieństwa, lecz jako wymóg spełniania tych aksjomatów i twierdzeń tego rachunku, które są szczególnie oczywiste. Takie rozluźnienie kryterium dopuszczalności przekształca je w kryterium nieostre, rodzi bowiem pytanie o to, które aksjomaty czy twierdzenia są „szczególnie oczywiste". W niektórych wypadkach odpowiedź wydaje się być prosta, np. aksjomaty nieujemności i unormowania są „szczególnie oczywiste", w innych już nie, np. nie jest jasne, czy aksjomat przeliczalnej addytywności lub twierdzenie Bayesa są „szczególnie oczywiste" czy nie. Nie bardzo jednak wiadomo, jak tej nieostrości można by uniknąć.

Podsumowując, kryteria dopuszczalności, stwierdzalności, stosowalności, jasności i uniwersalności stanowią podstawę oceny interpretacji prawdopodobieństwa, nie przesądzają jednak w jednoznaczny sposób o ich akceptacji lub odrzuceniu. Ocena danej interpretacji ma więc charakter swoistego „ważenia" różnych kryteriów, którego ostateczny wynik będzie w dużej mierze zależał od preferencji filozoficznych oceniającego. Jak słusznie pisze Hájek, można przyjąć nawet interpretację niedopuszczalną (ale, jak sądzę, nie *rażąco* niedopuszczalną!), jeśli dobrze spełnia pozostałe kryteria[19]. Wydaje mi się, że można co najwyżej powiedzieć, że szczególnie rażące naruszenie któregoś z tych kryteriów, a zwłaszcza kryterium dopuszczalności lub jasności, dyskwalifikuje daną interpretację prawdopodobieństwa.

Celem niniejszej pracy jest ocena jednego ze stanowisk sformułowanych w toku filozoficznej dyskusji na temat tego, czym jest prawdopodobieństwo: stanowiska „skłonnościowego". Prezentację tego stanowiska należy jednak, jak sądzę, poprzedzić omówieniem innych stanowisk. Bez takiego omówienia trudno byłoby w pełni zrozumieć specyfikę stanowiska „skłonnościowego" i dokonać jego oceny. Owe stanowiska to interpretacje klasyczna, logiczna, subiektywna i częstościowa.

1.3. Interpretacja klasyczna

1.3.1. Prezentacja

Zakładając, że zbiór zdarzeń elementarnych Ω jest skończony, a same zdarzenia elementarne są równo-możliwe, można udowodnić na podstawie aksjomatów 1-3 i własności 1-7 prawdopodobieństwa, że dla każdego $A \subset \Omega$:

$$P(A) = K(A)/K(\Omega).$$

W tym wzorze $K(A)$ oznacza liczbę elementów zbioru A (kardynalność zbioru A), a $K(\Omega)$ liczbę elementów zbioru Ω (kardynalność Ω).

Dowód. Gdy $\Omega = \{\omega_1, \dots, \omega_n\}$ i oznaczymy $P(\{\omega_i\}) = x$ dla $i = 1, \dots, n$, wówczas

[19] Por. Hájek 2003a.

$$1 = P(\Omega) = \sum_{i=1}^{n} P(\{\omega_i\}) = nx,$$

zatem $x = 1/n$ i

$$P(A) = \sum_{(i:\ \omega_i \in A)} P(\{\omega_i\}) = K(A)/n = K(A)/K(\Omega).$$

Powyższy wzór nazywany jest „klasyczną definicją prawdopodobieństwa”[20]. Na gruncie tej definicji $P(A)$, czyli prawdopodobieństwo absolutne zdarzenia A, interpretuje się więc jako stosunek liczby zdarzeń elementarnych „sprzyjających” temu zdarzeniu do liczby wszystkich zdarzeń elementarnych. Odpowiednio, $P(A|B)$, czyli prawdopodobieństwo warunkowe zdarzenia A, interpretuje się jako stosunek liczby zdarzeń elementarnych równocześnie „sprzyjających” A i B do liczby zdarzeń elementarnych „sprzyjających” B (np. jeśli A oznacza zdarzenie „wypadnie liczba trzy w rzucie kostką”, a B oznacza zdarzenie „wypadnie liczba nieparzysta w rzucie kostką”, to $P(A|B) = 1/3$). W obu wypadkach zakłada się, że zdarzenia elementarne są równo-możliwe.

Definicję klasyczną przyjmowali mniej lub bardziej świadomie myśliciele z kręgu *Port-Royal*, Jacques Bernoulli i Thomas Bayes, a najdojrzalszą postać przyjęła ona w pracy Pierre’a S. Laplace’a z 1814 roku *Essai philosophique sur les probabilités*.

1.3.2. Krytyka

Klasyczne ujęcie prawdopodobieństwa nie jest kontrowersyjne w takim zakresie, w jakim mówiąc o nim, ma się na myśli wyłącznie powyższy wzór – w sposób oczywisty wynikający z przyjętych aksjomatów – na obliczanie wartości prawdopodobieństw dla pewnych bardzo prostych sytuacji. Chodzi o takie

---

[20] Por. Jakubowski, Sztencel 2002, s. 26.

sytuacje, w których zbiór zdarzeń elementarnych jest skończony, a same zdarzenia elementarne są równo-możliwe. Przestrzeń probabilistyczna odpowiadająca takim sytuacjom jest szczególnym przypadkiem Kołmogorowskiej przestrzeni probabilistycznej ze skończoną addytywnością. *Ujęcie to jest jednak wysoce kontrowersyjne, jeśli traktuje się je nie tylko jako wzór na obliczanie prawdopodobieństw dla wspomnianych sytuacji, lecz także jako definicję prawdopodobieństwa.* Przejdę do omówienia głównych zarzutów, jakie wysunięto przeciw temu drugiemu rozumieniu ujęcia klasycznego, tj. jako definicji prawdopodobieństwa[21].

Po pierwsze, definicja klasyczna prawdopodobieństwa ma ograniczony zakres zastosowania. Może być stosowana tylko w tych szczególnych sytuacjach, w których, po pierwsze, zdarzenia elementarne są równo-możliwe, oraz, po drugie, zdarzenia te dają się łatwo wyróżnić. Jednak z sytuacjami spełniającymi oba warunki równocześnie mamy do czynienia stosunkowo rzadko (są nimi przede wszystkim gry losowe, takie jak karty, kości, ruletka)[22]. Warto także podkreślić, że ów ograniczony zakres za-

[21] Interpretacja klasyczna została poddana krytyce już w XIX weku. Pierwszymi jej krytykami byli: L. Ellis, J. Venn i T. Fechner.

[22] Np. nie są równo-możliwe wyniki rzutów monetą obciążoną (prawdopodobieństwo wypadnięcia orła, jak i prawdopodobieństwo wypadnięcia reszki jest różne od 0.5), wyniki eksperymentu polegającego na oczekiwaniu, czy w następnej sekundzie próbka radioaktywnego pierwiastka wyemituje cząstkę alfa czy nie (gdy przyjmiemy np., że prawdopodobieństwo emisji wynosi 0.4), wyniki doświadczenia losowego, jakim są narodziny noworodka (płeć męska jest „bardziej możliwa"). Nie da się na gruncie tej definicji także określić prawdopodobieństwa np. tego, że mający 40 lat Jan Kowalski dożyje przyszłego roku, nie wiadomo bowiem, w jaki sposób można by ją zastosować do takiej sytuacji (zdarzenie „Czterdziestoletni Jan Kowalski dożyje 41 lat" nie jest przecież równie możliwe jak np. zdarzenie „Osiemdziesięcioletni Jan Kowalski dożyje 81 lat", gdyż prawdopodobieństwo przeżycia każdego następnego roku maleje wraz z wiekiem). Rzecz jasna, uparty zwolennik interpretacji klasycznej mógłby wciąż utrzymywać, że wszystkie sytuacje, w których mamy do czynienia z wynikami, które nie są *prima facie* równo-możliwe, dają się teoretycznie zredukować do odpowiednich przestrzeni równo-możliwych zdarzeń elementarnych (np. mógłby

stosowania definicji klasycznej wynika nie tylko stąd, że można ją stosować wyłącznie do doświadczeń losowych, których wyniki są równo-możliwe i łatwo wyróżnialne, ale także – dodatkowo – stąd, że można ją bezpośrednio stosować wyłącznie do takich doświadczeń losowych, których przestrzeń zdarzeń elementarnych jest skończona[23].

Po drugie, definicja klasyczna jest kolista. Łatwo ową kolistość wykryć, stawiając pytanie o sens wyrażenia „zdarzenia równo-możliwe", występującego w tej definicji. Wyrażenie to – jak podnosiło wielu krytyków tej interpretacji – znaczy po prostu „zdarzenia równo-prawdopodobne". Wynika stąd, że prawdopodobieństwo definiuje się tutaj za pomocą prawdopodobieństwa. Czy może uniknąć tej kolistości? Jak przypomina Rudolf Carnap, zwolennicy interpretacji klasycznej nie definiowali wyrażenia „równo-możliwy" przez wyrażenie „równo-prawdopodobny", lecz przez odwołanie się do zasady racji niedostatecznej (dziś – za Johnem M. Keynesem – nazywanej „zasadą indyferencji")[24]. Zasada ta mówi, że jeżeli nie mamy żadnych racji, które skłania-

---

twierdzić, że w przypadku rzutów kostką obciążoną na rzecz orła dla każdego rzutu istnieje odpowiednia przestrzeń równo-możliwych zdarzeń elementarnych, z których więcej niż połowa „sprzyja" wynikowi orzeł). Takie przekonanie zwolennika interpretacji klasycznej byłoby jednak niczym nie uzasadnioną – choć skądinąd interesującą – tezą metafizyczną. W każdym razie, nawet gdyby teza ta była prawdziwa, zarzut ograniczonej stosowalności nadal pozostawałby w mocy, trudno bowiem uwierzyć, że zwolennik interpretacji klasycznej byłby kiedykolwiek w stanie podać dokładne charakterystyki odpowiednich przestrzeni równo-możliwych zdarzeń elementarnych dla wszystkich wyników, które *prima facie* nie są równo-możliwe.

[23] Ażeby definicję klasyczną można było zastosować do przestrzeni nieskończonej (przeliczalnej lub nieprzeliczalnej), należy przekształcić tę przestrzeń w przestrzeń skończoną (np. w ten sposób, iż połączy się jej punkty w większe zbiory, których liczba jest skończona, a następnie założy się, że z racji istnienia symetrycznie zbilansowanych świadectw na rzecz tych zbiorów mają one równe prawdopodobieństwo). Kwestię możliwości zastosowania definicji klasycznej do przestrzeni nieskończonych analizuje szczegółowo Hájek w Hájek 2003a; por. także Hájek 2001, s. 365 i Hájek 2006, s. 11.

[24] Por. Carnap 2000, s. 31.

łyby nas przyjęcia, że zajdzie raczej jedno zdarzenie niż drugie, powinnyśmy uznać, że są one równo-możliwe. *Prima facie*, zasada ta uchyla zarzut kolistości. Jest to jednak wrażenie złudne. Jak zauważa Hájek, zasada indyferencji może być zastosowana w dwóch sytuacjach: (a) gdy nie mamy w ogóle żadnych świadectw empirycznych; lub (b) gdy mamy świadectwa empiryczne, które są symetrycznie zrównoważone (np. gdy wnioskujemy na temat prawdopodobieństw w oparciu o fizyczną symetrię obecną w doświadczeniu losowym). Wprawdzie w sytuacjach pierwszego rodzaju faktycznie udaje się uniknąć kolistości, jednak, jak podkreśla Hájek, sytuacje tego rodzaju są bardzo rzadkie. W sytuacjach drugiego rodzaju nie udaje się już jednak uniknąć kolistości. Przykładem takiej sytuacji jest rzut kostką. Przyjmujemy tutaj, że wypadnięcie ścianki z numerem pięć jest równie możliwe jak wypadnięcie każdej innej z pozostałych pięciu ścianek, gdyż mamy wiele świadectw przemawiających za takim wnioskiem, tj. świadectw symetrycznie zrównoważonych (jednorodna, sześcienna kostka do gry jest symetryczna, ma dziewięć płaszczyzn symetrii, znamy wyniki poprzednich doświadczeń i relacje innych osób na temat rzutów tą kostką, itd.). Kolistość ujawnia się jednak ponownie, gdy tylko próbujemy wyeksplikować pojęcie świadectw symetrycznie zrównoważonych. Hájek twierdzi przekonująco, że najbardziej naturalną eksplikacją tego pojęcia jest eksplikacja oparta na pojęciu równości prawdopodobieństw warunkowych: przy danych świadectwach E i możliwych wynikach $A_1, A_2, \ldots, A_n$ świadectwa są symetrycznie zrównoważone wtedy i tylko wtedy, gdy $P(A_1|E) = P(A_2|E) = \ldots\ P(A_n|E)$. Znaczy to jednak, że chcąc posłużyć się pojęciem świadectw symetrycznie zrównoważonych, musimy dokonać czynności ważenia tych świadectw, która wymaga przeprowadzenia przynajmniej domyślnych oszacowań probabilistycznych. Innymi słowy, zgodnie z tą eksplikacją możemy powiedzieć, że nie mamy żadnych racji, aby przyjąć, że wypadnie raczej ten wynik niż inny tylko wtedy, gdy wiemy, że zdarzenia te są równo-prawdopodobne.

Kolistość pojawia się więc ponownie. Warto jeszcze w tym kontekście zwrócić uwagę na rzecz następującą: wydaje się, że wartości prawdopodobieństw ustalone (zgodnie z interpretacją klasyczną) w sytuacji (a) są miarą naszych przekonań, gdyż opierają się na braku naszej wiedzy, natomiast wartości prawdopodobieństw ustalone (zgodnie z interpretacją klasyczną) w sytuacji (b) nie są już miarą naszych przekonań, lecz raczej miarą określonych cech świata, gdyż oparte są na świadectwach empirycznych dotyczących układu losowego. Wydaje się więc, że w klasycznym ujęciu prawdopodobieństwa „mieszczą się" dwa jego różne rozumienia – jako miary własności naszych przekonań i jako miary własności świata – odpowiadające owym dwóm różnym sytuacjom, w których można zastosować zasadę indyferencji[25].

Po trzecie, zasada indyferencji generuje w pewnych sytuacjach tzw. paradoksy Bertranda[26]. Paradoksy Bertranda pokazują, że zasada indyferencji prowadzi do przypisania temu samemu zdarzeniu różnych wartości prawdopodobieństwa, jeśli do jego opisu zastosuje się różne parametryzacje powiązane definicyjnie w sposób nieliniowy (przy czym wszystkie te parametryzacje są równoważne). Oto przykład takiego paradoksu[27]. Fabryka produkuje sześciany o długości boku w przedziale [0, 1] metra. Postawmy następujące pytanie: jakie jest prawdopodobieństwo, że długość boku losowo wybranego sześcianu będzie w przedziale [0, 1/2] metra? Nie mamy żadnych racji, aby przyjąć, że długość boku będzie mieć wartość z przedziału [0, 1/2], a nie z przedziału [1/2, 1] – i *vice versa*, dlatego zgodnie z zasadą indyferencji zdarzeniu „długość boku sześcianu znajdzie się w przedziale [0, 1/2]" należy przypisać wartość 1/2. Poprzednie

---

[25] To ciekawe spostrzeżenie pochodzi od Michaela Strevensa; por. Strevens 1998.

[26] Od nazwiska XIX-wiecznego matematyka francuskiego Josepha Bertranda.

[27] Przykład ten pochodzi od Hájeka; por. Hájek 2003a. O paradoksie tym piszą także m.in. Earman, Salmon 1999, s. 74-77, Carnap 2000, s. 30-32, Gillies 2003, s. 14-24, i von Mises 1957, s. 66-77.

pytanie możemy jednak sformułować w inny – ale równoważny – sposób, opierając się na innej parametryzacji interesującego nas zdarzenia, powiązanej z poprzednią w sposób nieliniowy. Parametryzacja jest następująca: fabryka produkuje sześciany o powierzchni ścianek w przedziale [0, 1] metra kwadratowego. Wcześniejsze pytanie będzie teraz brzmieć tak: jakie jest prawdopodobieństwo, że powierzchnia boku losowo wybranego sześcianu będzie w przedziale [0, 1/4] metra kwadratowego? Tym razem odpowiedzią wynikającą z zastosowania zasady indyferencji będzie 1/4. Otrzymaliśmy więc dwie różne wartości prawdopodobieństw jako odpowiedzi na jedno pytanie (postawione na dwa *równoważne* sposoby). Jest to sytuacja kłopotliwa dla zwolennika interpretacji klasycznej, gdyż, jak pisałem w sekcji 1.2.2 niniejszego rozdziału, akceptowalna interpretacja prawdopodobieństwa musi spełniać warunek dopuszczalności, tj. musi być zgodna z aksjomatami rachunku prawdopodobieństwa, a aksjomaty te implikują, że prawdopodobieństwo jest funkcją i – w konsekwencji – każdemu zdarzeniu przypisuje tylko *jedną* wartość. Zauważmy zresztą, że w prezentowanym przykładzie zasada indyferencji prowadzi w istocie do przypisania danemu zdarzeniu nie dwóch różnych wartości prawdopodobieństw, lecz ich nieskończonej liczby. Wynika to stąd, że nasze wyjściowe pytanie możemy sformułować na nieskończenie wiele różnych sposobów w zależności od tego, jaką parametryzację interesującego nas zdarzenia przyjmiemy[28]. Podsumowując, paradoksy Bertranda pokazują, iż zasada indyferencji może prowadzić do niejednoznacznych wyników, co stanowi naruszenie kryterium dopuszczalności.

---

[28] Możemy powiedzieć np., że fabryka produkuje sześciany o objętości w przedziale [0, 1] metra sześciennego. Nasze pytanie brzmi wtedy: jakie jest prawdopodobieństwo, że losowo wybrany sześcian ma objętość w przedziale [0, 1/8]? Intuicyjną odpowiedzią będzie 1/8. W kolejnych sformułowaniach naszego pytania możemy mówić o kolejnych potęgach boku sześcianu.

Po czwarte, zasada indyferencji może prowadzić do jednoznacznych, ale nieintuicyjnych wyników. Laplace analizuje następujący przykład[29]. Pewną osobę informujemy zgodnie z prawdą, że moneta jest obciążona, nie mówimy jej jednak, czy chodzi o obciążenie na korzyść orła czy reszki. Następnie pytamy ją, jakie jest prawdopodobieństwo wyrzucenia reszki taką monetą. Jeśli osoba ta przyjmuje zasadę indyferencji, odpowie, że prawdopodobieństwo to wynosi 1/2, gdyż nie ma ona żadnych racji, aby przypisać wyższe prawdopodobieństwo któremuś z dwóch możliwych wyników. Laplace uważa, że taka odpowiedź jest właściwa. Nie jest to przekonujące. Wydaje się, że osoba ta powinna odpowiedzieć, że prawdopodobieństwo wyrzucenia reszki jest różne od 1/2.

Na zakończenie tej sekcji chciałbym się odnieść do głoszonej niekiedy tezy, że interpretacja klasyczna *zakłada*, iż prawdopodobieństwo nie jest miarą własności świata, lecz miarą naszej niewiedzy na temat świata, czyli miarą naszych przekonań. Ta teza, ściśle rzecz biorąc, jest nieprawdziwa. Definicja klasyczna nie zakłada, że prawdopodobieństwo jest przejawem naszej niewiedzy, nie zakłada bowiem nic na temat tego, do jakiego rodzaju obiektów odnosi się prawdopodobieństwo. Jak już wspominałem wcześniej, można co najwyżej powiedzieć, o ile jako element definicji klasycznej traktuje się zasadę indyferencji, że w definicji klasycznej „mieszczą się” dwa różne rozumienia prawdopodobieństwa – jako miary własności świata i jako miary naszych przekonań – odpowiadające dwóm różnym sytuacjom, w których można zastosować zasadę indyferencji. Jest natomiast prawdą, że zwolennicy tej definicji, np. Laplace, traktowali prawdopodobieństwo jako przejaw naszej niewiedzy – substytut rzetelnej wiedzy naukowej, a nie jako miarę własności świata[30]. Twierdzili,

---

[29] Por. Gillies 2003, s. 21.

[30] Gillies zwraca jednak uwagę na pewną niekonsekwencję w poglądach Laplace'a. Tytuł rozdziału VII dzieła Laplace'a *Essai philosophique sur les probabilités* brzmi „Des inégalités inconnues qui peuvent exister entre

że gdyby nasza wiedza naukowa była pełna, teoria prawdopodobieństwa byłaby zbędna – nie byłoby potrzeby odwoływania się do pojęcia prawdopodobieństwa. Taki pogląd na naturę prawdopodobieństwa nie jest jednak konsekwencją definicji klasycznej, lecz przejawem oświeceniowej wiary jej zwolenników w powszechny determinizm, motywowanej fascynacją osiągnięciami Newtonowskiej mechaniki[31].

1.3.3. Próba oceny

Nie jest do końca jasne, czy interpretacja klasyczna spełnia kryterium dopuszczalności: z jednej strony wydaje się, że spełnia (choć tylko ze skończoną addytywnością), gdyż opiera się na formule wynikającej w sposób oczywisty z aksjomatów Kołmogorowa; z drugiej jednak strony paradoksy Bertranda pokazują, że w pewnych sytuacjach zasada indyferencji – stanowiąca ważny składnik tej interpretacji – może prowadzić do przypisania jednemu zdarzeniu wielu różnych wartości prawdopodobieństwa, i tym samym do naruszenia aksjomatyki Kołmogorowa (która zakłada, że prawdopodobieństwo jest funkcją i – w konsekwencji – każdemu zdarzeniu przypisuje tylko jedną wartość). Interpretację

---

les chances que l'on supposes égales", a w swoim matematycznym dziele na temat prawdopodobieństwa *Théorie analytique des probabilités* z 1812 roku rozważa sytuację, w której szansa wypadnięcia orła wynosi $(1 + \alpha)/2$, a szansa wypadnięcia reszki $(1 - \alpha)/2$. Zdaniem Gilliesa zdaje się to świadczyć o tym, że obu wypadkach Laplace zakłada, iż istnieje obiektywne, nieznane prawdopodobieństwo uzyskania określonego wyniku. Stoi to w sprzeczności z jego poglądem, że prawdopodobieństwo jest tylko miarą naszej niewiedzy; jak pisze Gillies, „wygląda to tak, jak gdyby Laplace, rozwijając swoją teorię matematyczną, zapomniał o przyjętych przez siebie założeniach filozoficznych (Gillies 2003, s. 18)".

[31] Demon Laplace'a – najdoskonalsza personifikacja owej oświeceniowej wiary – nie musiałby odwoływać się do rachunku prawdopodobieństwa, gdyż znając wszystkie prawa natury i potrafiąc określić z dowolną dokładnością warunki początkowe każdego zdarzenia, potrafiłby jednoznacznie przewidzieć, czy dane zdarzenie nastąpi czy nie; ludzie są jednak skazani na niewiedzę i dlatego muszą posługiwać się pojęciem prawdopodobieństwa.

klasyczną można uznać za stwierdzalną i stosowalną tylko w tych sytuacjach, w których zasada indyferencji prowadzi do jednoznacznych wyników (takimi sytuacjami są np. rzuty monetą, rzuty kostką, itd.). Zakres zastosowania interpretacji klasycznej jest bardzo ograniczony. Wydaje się więc, że interpretacja klasyczna rażąco narusza kryterium uniwersalności. Podsumowując, interpretację klasyczną należy uznać za nieakceptowalną, gdyż z uwagi na zasadę indyferencji ma poważne problemy z kryterium dopuszczalności, rażąco narusza kryterium uniwersalności, oraz nie implikuje wyraźnego stanowiska w kwestii tego, czym jest prawdopodobieństwo (czy miarą własności świata, czy naszych przekonań).

1.4. Interpretacja logiczna

1.4.1. Prezentacja

Centralną tezę interpretacji logicznej można wysłowić na dwa sposoby[32]. Po pierwsze, można powiedzieć, że na gruncie tej interpretacji prawdopodobieństwo jest funkcją przekonań, nie dowolnej, lecz doskonale racjonalnej osoby co do prawdziwości zdania A względem zdania B (lub zbioru zdań). Po drugie, można powiedzieć, że na gruncie tej interpretacji prawdopodobieństwo jest funkcją określającą stopień, w jakim jakieś zdanie B (lub zbiór zdań) implikuje zdanie A[33]. W pierwszym sformułowaniu, prawdopodobieństwo dotyczy przekonań doskonale racjonalnej

[32] Najbardziej systematyczną wersję tej interpretacji przedstawił R. Carnap. Wcześniejsi zwolennicy interpretacji logicznej to m.in. W. E. Johnson, H. Jeffreys, J. Łukasiewicz, J. M. Keynes, W. Kneale, F. Waismann, L. Wittgenstein (por. np. tezy 4.464, 5.15-5.156 *Tractatus Logico-Philosophicus*). Za prekursorów tego ujęcia uchodzą G. W. Leibniz i B. Bolzano.

[33] Zdanie A może wynikać logicznie ze zdania B – mamy wtedy $P(A|B) = 1$; może częściowo wynikać logicznie – mamy wtedy $0 < P(A|B) < 1$, lub może wynikać w stopniu zerowym – mamy wtedy $P(A|B) = 0$.

osoby, w drugim – logicznej relacji między zdaniami A i B. Oba sformułowania są, jak się zdaje, równoważne: jeśli zdanie B częściowo (tzn. w stopniu p, gdzie $p < 1$) implikuje zdanie A, to jest rzeczą racjonalną wierzyć w prawdziwość A w stopniu p, gdy uznaje się za prawdziwe zdanie B. Drugie sformułowanie wydaje się być jednak bardziej adekwatne niż pierwsze, gdyż jest względem niego pierwotne (podmiot jest doskonale racjonalny, *ponieważ* określa prawdopodobieństwo zdania A kierując się względami czysto logicznymi) oraz jest nie-metaforyczne (pojęcie podmiotu doskonale racjonalnego jest użyteczną metaforą).

Z podstawową tezą interpretacji logicznej, iż prawdopodobieństwo zdania A na gruncie zdania B jest miarą stopnia, w jakim zdanie B pociąga za sobą zdanie A, wiążą się – mniej lub bardziej ściśle – tezy bardziej szczegółowe (należy podkreślić, że nie każdy ze zwolenników tej interpretacji byłby gotów podpisać się pod każdą z nich). Przedstawię najważniejsze z tych tez. Po pierwsze, zdania probabilistyczne nie są zdaniami faktualnymi, lecz logicznymi, a rachunek prawdopodobieństwa jest uogólnieniem logiki dedukcyjnej na sytuacje, w których między zdaniami zachodzi stosunek częściowego wynikania logicznego[34]. Po drugie, zdania probabilistyczne – jako dotyczące stosunków treściowych między zdaniami – są analityczne[35]. Po trzecie, po-

---

[34] Jak zobaczymy w dalszej części tego rozdziału, w odróżnieniu od prawdopodobieństw logicznych prawdopodobieństwa częstościowe i skłonnościowe nie są charakterystyką stosunku logicznego między zdaniami, lecz związku empirycznego między zdarzeniami, co znaczy, że w tych ujęciach zdania probabilistyczne są zdaniami faktualnymi, a nie logicznymi.

[35] Ale np. W. Kneale twierdził, że są one syntetyczne *a priori*. Zresztą, różnice w ujęciach interpretacji logicznej przyjmowanych przez poszczególnych filozofów są dość liczne. Nie będę próbował wyliczać ich wszystkich; wskażę tylko na kilka różnic między ujęciami Keynesa i Carnapa. Keynes bronił poglądu, że stosunkom logicznym między zdaniami nie zawsze można przypisać konkretną wartość liczbową. Twierdził także, że istnieją takie prawdopodobieństwa, które są nieporównywalne, tj. takie, o których nie można powiedzieć, że jedno jest z nich wyższe (lub równe) od drugiego (innymi słowy, zdaniem Keynesa, może zdarzyć się tak, iż potrafimy mówić o praw-

nieważ interpretacja logiczna zakłada, iż prawdopodobieństwo jest miarą relacji logicznej między zdaniami, można powiedzieć, że podstawowym typem prawdopodobieństw, jaki pojawia się na gruncie tej interpretacji, są prawdopodobieństwa warunkowe $P(A|B)$[36]. Prawdopodobieństwa absolutne $P(A)$ są więc w ramach tej interpretacji traktowane co do zasady jako eliptyczny zapis prawdopodobieństw warunkowych. Owe eliptyczne prawdopodobieństwa warunkowe mogą być dwojakiego rodzaju: (i) zdanie *A*, któremu przypisujemy prawdopodobieństwo, może być zrelatywizowane do jakiegoś – przyjętego *implicite* – standardowego zbioru zdań; (ii) zdanie to może być zrelatywizowane do pustego zbioru świadectw, tj. tylko do zbioru prawd logicznych[37].

---

dopodobieństwie zdania *A* względem zdania *B* i prawdopodobieństwie zdania *C* względem zdania *D*, ale nie potrafimy powiedzieć, czy to pierwsze prawdopodobieństwo jest wyższe od drugiego czy nie). Chodzi tutaj – zdaniem Keynesa – nie tylko o kwestie epistemologiczne, ale także o to, że stosunek logiczny miedzy zdaniami może nie mieć konkretnej wartości liczbowej, a prawdopodobieństwa mogą być co do zasady nieporównywalne. Carnap odrzucał te tezy. Odrzucał także pogląd Keynesa, że do przypisania prawdopodobieństwa danemu zdaniu dochodzi się na drodze intuicji (Keynes twierdził, że są pewne zdania, o których wiemy bezpośrednio, że są prawdziwe, i od których zależą wszystkie racjonalne sądy probabilistyczne). Odpowiedzią Carnapa na ten pogląd była jego koncepcja prawdopodobieństw jako stopni konfirmacji. O ile więc dla Keynesa prawdopodobieństwo było nieanalizowalną relacją logiczną, o tyle dla Carnapa było relacją analizowalną; por. Keynes 1963 i Carnap 1950. Omówienie tych różnic można znaleźć także np. w Nekrasas 1992, s. 99-100 i Gillies 2003, s. 30.

[36] Relatywizacja prawdopodobieństw nie jest jednak jakąś *differentia specifica* interpretacji logicznej, gdyż, jak już wspominałem w sekcji 1 niniejszego rozdziału, wszystkie interpretacje w jakiś sposób relatywizują prawdopodobieństwa.

[37] Pojawia się pytanie, jaki jest stopień wynikania *A* w tym drugim wypadku. Nie może to być stopień równy zero, gdyż taki sam musiałby być stopień wynikania negacji *A*, podczas gdy jedno z tych zdań musi być prawdziwe. Bardziej intuicyjne jest przyjęcie założenia, że pusty zbiór świadectw implikuje oba zdania w identycznym stopniu równym 1/2. Takie przypisanie prawdopodobieństw jest oczywiście oparte na zasadzie indyferencji. Także ono ma jednak niezbyt intuicyjną konsekwencję. Wyobraźmy sobie, że dys-

Należy jednak podkreślić, iż nie jest tak, że interpretacja logiczna całkowicie wyklucza możliwość stwierdzenia o jakimś zdaniu, że jest ono „samo w sobie”, tzn. bezwarunkowo, prawdopodobne w takim czy innym stopniu. Prawdopodobieństwa absolutne – tzw. prawdopodobieństwa logiczne *a priori* – przypisywane są na gruncie tej interpretacji – przynajmniej w ujęciu Carnapa – możliwym stanom rzeczy[38]; prawdopodobieństwa te są niezbędne, aby móc następnie określać wartości prawdopodobieństw „właściwych”, tj. warunkowych. Warto jeszcze dodać, że owo przypisywanie prawdopodobieństw absolutnych możliwym stanom rzeczy dokonuje się za pośrednictwem zasady indyferencji. Po czwarte, uznanie prawdopodobieństwa za charakterystykę relacji między zdaniami umożliwia nadanie mu obiektywnego sensu nawet w świecie deterministycznym (tę kwestię omawiam dokładnie w sekcji 1.8 niniejszego rozdziału). Po piąte, jeśli rachunek prawdopodobieństwa jest uogólnieniem logiki dedukcyjnej na sytuacje, w których między zdaniami zachodzi stosu-

---

ponujemy bardzo słabymi świadectwami empirycznymi B na rzecz zdania A. Otóż w świetle tych świadectw wartość $P(A|B)$ jest na pewno mniejsza od 1/2, czyli mniejsza niż wtedy, gdy nie mamy żadnych świadectw na rzecz A.

[38] W swoim fundamentalnym dziele *Logical Foundations of Probability* Carnap wyróżnił klasę bardzo prostych języków formalnych, składających się ze skończonej liczby logicznie niezależnych predykatów jednoargumentowych (oznaczających własności) stosowanych do przeliczalnie wielu nazw indywiduowych (oznaczających indywidua) oraz ze zwykłych spójników logicznych. Zdania, które w sposób możliwie szczegółowy na gruncie tego języka opisują wszystkie indywidua, Carnap nazywa „opisami stanu” („*state descriptions*”) (dany opis stanu jest więc opisem możliwego stanu rzeczy). Każde takie zdanie jest koniunkcją pełnych opisów każdego z indywiduów, przy czym każdy opis danego indywiduum jest koniunkcją, której argumentami są wszystkie predykaty tego języka (każdy z tych – występujących jednokrotnie – predykatów jest albo poprzedzony negacją, albo nie). Każde zdanie tego języka jest równoważne jakiejś alternatywie wzajemnie wykluczających się opisów stanu. Powyższa charakterystyka oparta jest na Hájek 2001, s. 365-366, Hájek 2003a, Grobler 2006, s. 38-45, i Mortimer 1982, s. 63-82. Oryginalne teksty Carnapa na ten temat to przede wszystkim – wspomniany – Carnap 1950, a także np. Carnap 1963.

nek częściowego wynikania logicznego, to można próbować go wykorzystać do budowy logiki indukcji, czyli ilościowej teorii konfirmacji hipotez naukowych przez świadectwa empiryczne, gdyż, jak się zdaje, w rozumowaniach indukcyjnych przesłanki, tj. świadectwa empiryczne, częściowo implikują wnioski, tj. hipotezy naukowe. Wartości prawdopodobieństwa są w tym kontekście „miarą" relacji logicznej między świadectwami empirycznymi i daną hipotezą naukową, czyli miarą stopnia, w jakim świadectwa empiryczne „wspierają", „potwierdzają", czy „konfirmują" tę hipotezę[39]. Najsłynniejszą próbę tego rodzaju podjął Carnap. Carnap początkowo przyjmował, że istnieje jedna właściwa miara potwierdzania zdań przez inne zdania, czyli funkcja konfirmacji[40]; z takiego założenia wynika, że jeśli osoby dosko-

[39] Prawdopodobieństwo wyjściowe – absolutne – hipotezy naukowej, tj. prawdopodobieństwo hipotezy naukowej przed uzyskaniem jakichkolwiek świadectw empirycznych, zależy od liczby możliwych stanów rzeczy, z jakimi hipoteza ta jest zgodna: im większa jest ta liczba, tym wyższe jest owo prawdopodobieństwo. Ponieważ taki sposób określania wartości prawdopodobieństw absolutnych hipotez naukowych rzadko daje się zastosować w praktyce, zwolennicy interpretacji logicznej określają te wartości w ten sposób, że stosują zasadę indyferencji bezpośrednio do hipotez naukowych (zasadę tę można w tym kontekście ująć tak: jeśli nie dysponuje się żadnymi świadectwami na poparcie którejś z n wzajemnie wykluczających się i dopełniających się hipotez, należy przyjąć, że prawdopodobieństwo każdej z nich jest równe $1/n$).

[40] Dowolna miara prawdopodobieństwa $m(-)$ określona na opisach stanu (jak już wspominałem, można te miarę interpretować jako prawdopodobieństwo absolutne, tj. logiczne *a priori*), o ile spełnia aksjomaty rachunku prawdopodobieństwa (tj., m.in., przypisuje każdemu ze stanów wartość z przedziału [0, 1] i wartości te sumują się do 1), wyznacza automatycznie – zgodnie z definicją prawdopodobieństwa warunkowego – funkcję konfirmacji $c(-,-)$: $c(H|E) = m(H \wedge E)/m(E)$. W ten sposób powstaje dla danego języka logika indukcji, która każdej z par zdań $< H, E >$ tego języka przypisuje jedną wartość prawdopodobieństwa, określającą stopień konfirmacji H przez E. Nawet dla bardzo prostych języków istnieje nieskończenie wiele możliwych miar m, i – w rezultacie – funkcji c. Carnap opowiadał się za miarą, którą zapisywał jako $m*$, i za odpowiadającą jej funkcją konfirmacji $c*$. Miara $m*$ przypisuje równe prawdopodobieństwa tzw. „opisom struktury" („*structure de-*

nale racjonalne różnią się w ocenie stopnia potwierdzenia danego zdania A, to przyczyną musi być to, że dysponują innymi zbiorami świadectw. Założenie to odrzucił w swoich późniejszych pracach, uznając za dopuszczalne wiele – ściśle, *continuum* – różnych funkcji konfirmacji.

1.4.2. Krytyka

Na pierwszy rzut oka logiczne ujęcie prawdopodobieństwa jest przekonywujące. Przyjmijmy np., że zdanie A brzmi „następny zaobserwowany kruk będzie czarny", zdanie B – „zaobserwowano jednego kruka, i okazał się czarny", zaś zdanie C – „zaobserwowano setki kruków, i wszystkie okazały się czarne". Otóż wydaje się, że można powiedzieć, iż $P(A|B) < P(A|C)$, a więc, że zdanie C w większym stopniu niż zdanie B uprawdopodabnia zdanie A, oraz że wartości $P(A|B)$ i $P(A|C)$ istnieją jako obiektywne fakty i dają się przynajmniej w przybliżony sposób określić. Interpretacja logiczna opiera się na intuicji, że tego rodzaju wnioski są poprawne. Niemniej, przeciwko tej interpretacji sformułowano szereg poważnych zarzutów. Omówię cztery z nich.

Po pierwsze, hipotezy probabilistyczne stanowią podstawę do formułowania empirycznych prognoz zjawisk, a empiryczna trafność tych prognoz byłaby niezrozumiała, gdyby jedyną interpretacją prawdopodobieństwa była interpretacja logiczna, która zakłada, że wartości prawdopodobieństwa nie są miarą możliwości

---

scriptions"), tj. maksymalnym zbiorom takich opisów stanu, z których każdy może być uzyskany z dowolnego innego opisu z tego zbioru przez permutację nazw indywiduowych (opisy struktury są więc klasami równoważności opisów stanów utworzonymi przez permutacje nazw). Prawdopodobieństwo danego opisu struktury jest następnie dzielone po równo między wszystkie elementy tego opisu, tj. między wszystkie konstytuujące opis struktury opisy stanu. Prostą konsekwencją tak zdefiniowanej miary jest to, że każdy homogeniczny opis stanu będzie miał wyższe prawdopodobieństwo niż opis heterogeniczny; por. Hájek 2001, s. 365-366, Hájek 2003a, Grobler 2006, s. 38-45, i Mortimer 1982, s. 63-82.

wystąpienia zjawisk, lecz miarą „siły", z jaką z danego zdania lub zbioru zdań wynikają inne zdania. Na przykład, kiedy mówimy, że prawdopodobieństwo tego, że jutro będzie padał deszcz wynosi 0.2, chcemy powiedzieć *coś o świecie zewnętrznym*, nie zaś o tym, że określony zbiór zdań (dane meteorologiczne, jakimi dysponujemy) pociąga za sobą z „siłą" 0.2 zdanie „jutro będzie padał deszcz"[41]. Prawdziwość wielu hipotez probabilistycznych jest więc empiryczna, a nie logiczna[42]. Zarzut ten pokazuje wyraźnie, że interpretacja logiczna nie jest interpretacją uniwersalną.

Po drugie, Frank Ramsey zakwestionował tezę Keynesa, że logiczne relacje prawdopodobieństwa między zdaniami mogą być stwierdzone na drodze intuicji. Zwrócił uwagę na fakt, że między osobami, którzy twierdzą, że potrafią dostrzec takie relacje, nie ma zgody co do tego, jakie konkretnie relacje łączą dwa dowolne zdania. W rezultacie doszedł do wniosku, że logiczna intuicja jest psychologiczną iluzją, której podlegamy dzięki zażyłości z danym przedmiotem, i która następnie rodzi w nas błędne przekonanie o istnieniu jakichś relacji logicznych między zdaniami[43]. Szczególnie uderzał go fakt, że nie potrafimy ocenić relacji logicznych między najprostszymi zdaniami (np. zdaniem

---

[41] Por. von Mises 1957, s. 96.

[42] Różnicę między interpretacją logiczną i jakąś interpretacją „empiryczną" (np. częstościową lub skłonnościową) dobrze ilustruje także następujący przykład Poppera. W formule $P(A|B) = r$, za A podstawiamy „a jest elipsą", za B „a jest orbitą planetarną". Można pokazać, że wartość r będzie zależeć od tego, jaką interpretację prawdopodobieństwa przyjmiemy: jeśli będzie to interpretacja logiczna, wtedy $r = 0$, ponieważ orbita może przybrać nieskończenie wiele kształtów; jeśli będzie to interpretacja częstościowa lub skłonnościowa, wtedy $r = 1$, gdyż zgodnie z pierwszym prawem Keplera każda orbita planetarna ma kształt elipsoidalny; por. Popper 1983, s. 288-290.

[43] Jak przypomina Gillies, intuicja logiczna okazuje się zwodnicza lub niejednoznaczna nawet w kontekście logiki dedukcyjnej, np. G. Frege przyjmował aksjomat, z którego B. Russell w prosty sposób wyprowadził swój słynny paradoks, D. Hilbert przyjmował prawo wyłączonego środka, które L. E. J. Brouwer odrzucał; por. Gillies 2003, s. 52-53.

„To jest czerwone” i zdaniem „Tamto jest niebieskie”)[44]. Opisany zarzut wydaje się trafny względem tych ujęć interpretacji logicznej (m.in. ujęcia Keynesa), które traktują relację logiczną między zdaniami jako nieanalizowalną – dostępną tylko dla intuicyjnego wglądu. Nie jest jednak w pełni trafny np. względem ujęcia Carnapa, który próbował dokonać analizy tej relacji.

Po trzecie, ujęcie Carnapa unika wprawdzie opisanego wyżej zarzutu Ramseya, napotyka jednak inny zarzut – arbitralności wyboru wartości prawdopodobieństw. Arbitralność ta pojawia się w kilku miejscach konstrukcji Carnapa. Po pierwsze, wartości prawdopodobieństw absolutnych przypisanych różnym możliwym stanom rzeczy (i w konsekwencji, wartości prawdopodobieństw warunkowych) zależą od wyboru języka, jakiego użyjemy do ich reprezentacji, oraz od wyboru funkcji konfirmacji. Po drugie, reguła „późnego” Carnapa określająca stopień indukcyjnego potwierdzenia zawiera parametr λ, który, nieco upraszczając, określa szybkość, z jaką uczymy się na podstawie doświadczenia. Otóż wartość tego parametru można wybierać dowolnie. Wybór wartości z jednego ekstremum pozwala wyprowadzać natychmiastowe wnioski indukcyjne (na podstawie obserwacji jednego czarnego kruka, możemy wnioskować, że kolejny też będzie czarny), zaś wybór wartości z drugiego ekstremum wyklucza możliwość wnioskowania, że kolejny kruk będzie czarny nawet po tym, jak zaobserwowano setki czarnych kruków[45]. Po trzecie, o arbitralności można mówić także w sytuacji wyboru świadectw, względem których określana jest wartość prawdopodobieństwa danej hipotezy[46]. Prawdopodobieństwo logiczne opisuje relację potwierdzania danego zdania A przez ja-

[44] Por. Ramsey 1990.

[45] Sam Carnap podał pozalogiczne argumenty za konkretną wartością parametru λ, dochodząc do reguły indukcyjnej równoważnej „regule sukcesji” Laplace’a, z której wynika, że jeśli i z n zaobserwowanych kruków były czarne, to zdaniu, że następny kruk będzie czarny, należy przypisać prawdopodobieństwo $(i + 1)/(n + 2)$.

[46] Na ten fakt zwraca uwagę m.in. A. Ayer w Ayer 1963, s. 188-202.

kieś inne zdanie *B* (lub zbiór zdań). Wartość tego prawdopodobieństwa zależy więc od wyboru określonego zbioru zdań z nieskończonej liczby takich zbiorów. Wobec powyższego pojawia się pytanie o to, według jakich kryteriów powinno się dokonywać takiego wyboru i tym samym wyboru wartości prawdopodobieństw. Wydaje się, że w ramach samej interpretacji logicznej nie sposób uzasadnić, że jeden wybór jest bardziej adekwatny niż jakiś inny[47]. Carnap twierdził, że powinniśmy przyjąć zasadę *total evidence*, zgodnie z którą nasze oszacowania prawdopodobieństw powinny być oparte na możliwie szerokim zbiorze świadectw. Jest to postulat zdroworozsądkowy i przekonujący, nie bardzo jednak wiadomo, dlaczego, jeśli przyjmujemy interpretacją logiczną *jako jedynie uprawnioną*, mielibyśmy uznać go za trafny[48]. Nie możemy wszak odwołać się do tego, że tak określone prawdopodobieństwo jest bliższe „rzeczywistemu" niż określone na podstawie węższego zbioru świadectw, gdyż zakładamy, że żadne rzeczywiste prawdopodobieństwo nie istnieje. Podsumowując, zwolennicy interpretacji logicznej mają trudności nie tylko ze wskazaniem jasnych kryteriów wyboru języka opisu, funkcji konfirmacji oraz wartości parametru dotyczącego „szybkości" indukcyjnego uczenia się, ale także ze wskazaniem, na którym z wielu możliwych zbiorze świadectw należy oprzeć oszacowania wartości prawdopodobieństw. W każdym razie, chcąc usunąć opisane niejednoznaczności, trzeba posłużyć się argumentami pozalogicznymi. Należy podkreślić, że Carnap doskonale zdawał sobie sprawę z powyższych problemów. Był w pełni świadom tego, że o ile generalnie istnieje konsens w kwestii tego, jakie zdania wynikają dedukcyjnie z in-

[47] Na gruncie interpretacji częstościowej ten zarzut wiąże się z problemem klasy odniesienia, o czym będzie mowa w sekcji 1.6.2 tego rozdziału.

[48] Gwoli jasności należy dodać, że Carnap nie uważał, że interpretacja logiczna jest jedyną uprawnioną interpretacją prawdopodobieństwa; dopuszczał także interpretację częstościową (prawdopodobieństwa określone na podstawie pierwszej nazywał prawdopodobieństwami1, zaś określone na podstawie drugiej – prawdopodobieństwami2); por. Carnap 1945.

nych zdań, o tyle istnieje rozbieżność opinii w kwestii tego, w jakim stopniu określone zdania wspierają indukcyjnie inne zdania. Zgadzał się, że istnieje wiele systemów logiki indukcyjnej, które – z czysto logicznego punktu widzenia – są równie dopuszczalne. Uważał jednak, że kierując się względami pozalogicznymi, można z tego zbioru logik wybrać taką, która będzie najlepiej służyć naszym praktycznym celom.

Po czwarte, warto wspomnieć o najważniejszym punkcie Popperowskiej krytyki interpretacji logicznej. Popper traktował prawdopodobieństwo logiczne jako uprawnioną interpretację formuły $P(A|B)$, krytykował jednak ideę probabilistycznej logiki indukcji, opartej na założeniu, że stopień konfirmacji hipotezy empirycznej można utożsamić z jej prawdopodobieństwem logicznym. Zdaniem Poppera, prawdopodobieństwo logiczne nie może być miarą wartości hipotez empirycznych i – w konsekwencji – miarą racjonalności naszych przekonań odnośnie tych hipotez. Oto jeden z jego najważniejszych argumentów na rzecz tej tezy. Jeśli mamy skończoną liczbę świadectw empirycznych B i zdanie ściśle uniwersalne A, mające potencjalnie nieskończoną liczbę egzemplifikacji, wtedy z uwagi na to, iż A „wykracza" nieskończenie poza B, $P(A|B)$ będzie równe 0. Funkcje konfirmacji przypisują więc zdaniom ogólnym w nieskończonym uniwersum zerową konfirmację. Gdyby więc prawdopodobieństwo logiczne było miarą racjonalności naszego przekonania co do A, stopień ten byłby zerowy, co wydaje się absurdalne, zważywszy na to, iż przecież na podstawie doświadczenia dowiadujemy się coraz więcej o uniwersalnych prawach. Popper konkluduje, że miarą, którą zwiększamy, podnosząc stopień racjonalności naszych przekonań co do A, nie może być prawdopodobieństwo logiczne, lecz jakaś inna miara, którą nazywa „stopniem koroboracji". Wysoki stopień koroboracji i – w konsekwencji – wysoki stopień prawdoupodobnienia (*verisimilitude*) uzyskuje odważna, tj. mało prawdopodobna hipoteza, która przeszła pomyślnie testy empiryczne. Podsumowując, zdaniem Poppera, nie prawdopodobieństwo logiczne $P(A|B)$, jak chciał Carnap, lecz stopień koroboracji, jest miarą racjonalności naszego przekonania odnośnie A w świetle B.

1.4.3. Próba oceny

Interpretacja logiczna jest dopuszczalna (w ramach interpretacji logicznej algebra Boole'a jest interpretowana w logice zdań, a nie w rachunku zbiorów, ale nie ma to żadnego znaczenia, gdyż, jak już wspominałem w sekcji 1.1 tego rozdziału, oba ujęcia są izomorficzne)[49]. Ponadto, w danym języku wartości funkcji konfirmacji są stwierdzalne. Nie jest całkiem jasne, czy interpretacja logiczna spełnia kryterium stosowalności. Wydaje się, że nie spełnia, gdyż stosunkowo rzadko można na jej podstawie wygenerować wartości prawdopodobieństw, które mogłyby być „przewodnikiem w życiu". Interpretacja logiczna nie spełnia kryterium uniwersalności: w praktyce naukowej i dyskursie codziennym formułuje się liczne wypowiedzi probabilistyczne, które wydają się odnosić do świata, a nie do relacji między zdaniami. Generalnie rzecz biorąc, interpretacja logiczna wydaje się być interpretacją akceptowalną. Innym problemem jest to, czy interpretacja logiczna może być podstawą budowy ilościowej teorii konfirmacji hipotez naukowych. Wśród filozofów nauki panuje duża rozbieżność zdań na ten temat: niektórzy stają po stronie Carnapa, inni Poppera. Spór ten pozostaje więc wciąż nierozstrzygnięty[50].

[49] Warto zauważyć, że interpretacja logiczna spełnia, choć pusto, także aksjomat przeliczalnej addytywności, gdyż w dowolnym skończonym języku nie może istnieć nieskończony ciąg wykluczających się parami zdań.

[50] Por. na ten temat np. Michalos 1971, Popper 1999, s. 471-493.