Przegląd Filozoficzny – Nowa Seria
R. 25: 2016, Nr 2 (98), ISSN 1230–1493

Epistemologia

Paweł Pruski

Problem wrażliwości miary Bayesjańskiej Teorii Potwierdzania – analiza Carnapowskiego argumentu z symetrii*

Słowa kluczowe: *Bayesjańska Teoria Potwierdzania, prawdopodobieństwo, symetria, R. Carnap*

Wprowadzenie

Według Bayesjańskiej Teorii Potwierdzania, zdarzenie E potwierdza hipotezę H wtedy i tylko wtedy, gdy H i E są ze sobą dodatnio skorelowane probabilistycznie. Istnieje wiele równoważnych, a zarazem nieredukowalnych sposobów wyrażenia korelacji pomiędzy H oraz E. Powyższa sytuacja prowadzi do problemu związanego z wyborem odpowiedniej miary potwierdzania. Istnieje silna zależność pomiędzy wyborem miary potwierdzania oraz efektywnością Bayesjańskiej Teorii Potwierdzania (jej podatnością na określone typy paradoksów). Powyższa korelacja określana jest jako tzw. problem wrażliwości miary[1]. Konstrukcja adekwatnej miary potwierdzania, czyli takiej, która pozwoli ominąć szereg problemów, jest jednym z głównych wyzwań, jakie stawiają przed sobą zwolennicy Bayesjańskiej Teorii Potwierdzania. Autorem jednej z kluczowych miar potwierdzania jest Rudolf Carnap.

---

* Praca powstała w wyniku realizacji projektu badawczego nr 2013/09/N/HS1/00902 finansowanego ze środków Narodowego Centrum Nauki.

[1] Po raz pierwszy tzw. problem wrażliwości miary sformułowany został przez B. Fitelsona w *Studies in Bayesian Confirmation Theory*, University of Wisconsin 2001.

Carnap określany jest mianem prekursora współczesnej wersji Bayesjańskiej Teorii Potwierdzania. Jego zdaniem, stosowanie w nauce procedury zmierzające do oceny relacji pomiędzy hipotezą a wynikami obserwacji prowadzą do potwierdzania lub odrzucania niektórych hipotez, a co za tym idzie, stałego rozwoju nauki. Celem teorii potwierdzania jest zdefiniowanie funkcji prawdopodobieństwa (funkcji miary), dzięki której możliwe jest przyporządkowanie hipotezom określonych stopni uzasadnienia ze względu na wyniki obserwacji. Rezultatem powyższych założeń jest obszerny system obejmujący formalizację wnioskowań indukcyjnych za pomocą rachunku prawdopodobieństwa, który przedstawiony został przez Carnapa w *Logical Foundations of Probability* (1962). Jednym z istotnych elementów powyższego systemu jest miara potwierdzania c, bazująca na szeregu dodatkowych warunków, jakie zostały na nią nałożone. Z perspektywy niniejszej pracy szczególnie interesujące są warunki dotyczące symetryczności miary potwierdzania.

Celem tekstu jest analiza carnapowskiej funkcji miary c w kontekście problemu wrażliwości miary Bayesjańskiej Teorii Potwierdzania. Rezultat analizy rzuca nowe światło na warunki nakładane na miarę potwierdzania, a w szczególności na własności związane z symetrycznością miary potwierdzania. Analiza pokazuje, że zbyt silna symetryczność prowadzi do niepożądanych konsekwencji.

W pierwszej części pracy przybliżona zostanie ogólna charakterystyka Bayesjańskiej Teorii Potwierdzania. Część druga skoncentrowana jest wokół tzw. problemu wrażliwości miary. Część trzecia obejmuje analizę carnapowskiej miary potwierdzania c. Ostatnia, czwarta część pracy zawiera podsumowanie zaproponowanej analizy.

1. Bayesjańska Teoria Potwierdzania

W obrębie współczesnej epistemologii oraz filozofii nauki teoria potwierdzania (*confirmation theory*) opiera się na zastosowaniu narzędzi logiki do analizy relacji pomiędzy danymi obserwacyjnymi oraz określoną hipotezą. Trzon problemu sprowadza się do odpowiedzi na pytanie – w jakim stopniu hipoteza jest potwierdzana przez dane obserwacyjne? Z perspektywy historycznej powyższa tematyka związana jest z problemem indukcji. Podobnie jak w przypadku prób formalizacji związków indukcyjnych (np. logika indukcji), tak również we współczesnych teoriach potwierdzania kluczową rolę odgrywa rachunek prawdopodobieństwa. Jedną z teorii potwierdzania posługujących się rachunkiem prawdopodobieństwa jest Bayesjańska Teoria Potwierdzania (BCT).

W świetle BCT potwierdzanie jest relacją pomiędzy zdarzeniami E (*evidences*), hipotezą H (*hypothesis*), dodatkowymi założeniami tła B (*background* *assumptions*), mierzoną za pomocą prawdopodobieństwa Pr. Stopień potwierdzenia hipotezy H przez zdarzenie E w stosunku do B definiowany jest jako prawdopodobieństwo warunkowe danej hipotezy ze względu na zdarzenie E oraz informacje tła B, $Pr(H \mid E \cap B)$.

Jednym z centralnych pojęć BCT jest tzw. miara potwierdzania (*confirmation measure*). Zgodnie z powyższą teorią miara potwierdzania powinna spełniać następujący warunek (W1). Funkcja rzeczywista $c: P \rightarrow \mathfrak{R}$ ze zbioru P przestrzeni probabilistycznej $\langle W, A, Pr \rangle$ jest miarą potwierdzania wtw., gdy dla każdej przestrzeni probabilistycznej $\langle W, A, Pr \rangle$ i dla każdego $H, E, B \in A$[2]:

$$c(H, E, B) > 0 \Leftrightarrow Pr(H \cap E|B) > Pr(H|B) \cdot Pr(E|B)$$
$$c(H, E, B) = 0 \Leftrightarrow Pr(H \cap E|B) = Pr(H|B) \cdot Pr(E|B)$$
$$c(H, E, B) < 0 \Leftrightarrow Pr(H \cap E|B) < Pr(H|B) \cdot Pr(E|B)$$

Oprócz wskazanego powyżej warunku (W1), jaki powinna spełniać miara potwierdzania, BCT nie nakłada żadnych dodatkowych ograniczeń. Konsekwencją jest możliwość stworzenia wielu nierównoważnych, alternatywnych miar stopnia, w jakim E i B potwierdzają H (zgodnie z Pr). W związku z mnogością możliwości, wielu autorów wprowadza dodatkowe warunki, tak aby zwiększyć oczekiwaną efektywność zaproponowanej miary potwierdzania. Obecnie w literaturze przedmiotu (por. Huber 2006; Fitelson 2010) wyróżnia się pięć podstawowych miar potwierdzania:

(m1) Carnapowska miara **c** (Carnap 1962)
$$\mathbf{c}(H, E \mid B) =_{df} Pr(H \cap E \mid B) - Pr(H \mid B) \cdot Pr(E \mid B)$$

(m2) Miara odległości **d** (Earman 1992)
$$\mathbf{d}(H, E \mid B) =_{df} Pr(H \mid E \cap B) - Pr(H \mid B)$$

(m3) Miara Gooda-Fitelsona **l** (Fitelson 1999, 2001; Good 1983)
$$\mathbf{l}(H, E \mid B) =_{df} \log [Pr(E \mid H \cap B) / Pr(E \mid -H \cap B)]$$

(m4) Miara Milne'a **r** (Milne 1996)
$$\mathbf{r}(H, E \mid B) =_{df} \log [Pr(H \mid E \cap B) / Pr(H \mid B)]$$

(m5) Miara Joyce'a-Christensena **s** (Christensen 1999, Joyce 1999).
$$\mathbf{s}(H, E \mid B) =_{df} Pr(H \mid E \cap B) - Pr(H \mid -E \cap B)$$

---

[2] Funkcja spełniająca powyższy warunek określana bywa również jako miara trafności (*relevance measure*).

Wybór którejś z powyższych miar niesie za sobą poważne konsekwencje natury teoretycznej oraz praktycznej. Z jednej strony związany jest z przyjęciem określonych założeń filozoficznych, z drugiej pozwala na zastosowanie jej w wybranym obszarze nauki. Z wyborem miary potwierdzania w związku z określonymi typami problemów łączy się tzw. problem wrażliwości miary.

2. Problem wrażliwości miary

Problem wrażliwości miary opiera się na relacji pomiędzy miarą potwierdzania oraz specyficzną grupą problemów dotyczących BCT. Wspomniane problemy obejmują wiele klasycznych paradoksów potwierdzania (np. paradoks kruków, paradoks ziembieskości) typowych dla probabilistycznych teorii potwierdzania. Obecnie wielu autorów analizuje wspomniane problemy w odniesieniu do zastosowania wybranych miar potwierdzania. Celem tych prób jest wyróżnienie typów miar potwierdzania, które pozwalają ominąć powyższe paradoksy. Do najbardziej znanych analiz miar potwierdzania (w obrębie klasycznych paradoksów potwierdzania) możemy zaliczyć:

(p1) „Paradoks przechodniości potwierdzenia" (Earman 1992),
(p2) „Paradoks ziembieskości" (Eells 1982),
(p3) „Paradoks ziembieskości" (Sober 1994),
(p4) „Paradoks kruków" (Horwich 1982).

Tym samym miary potwierdzania (m1) – (m5) posiadają różny stosunek do paradoksów (p1) – (p4). Obowiązywalność wskazanych paradoksów uzależniona jest od tego, która ze wskazanych miar potwierdzania została zastosowana. Innymi słowy, problem wrażliwości miary dotyczy relacji pomiędzy (m1) – (m5) a (p1) – (p4).

Bazując na obszernej analizie, jaką przeprowadził Fitelson (2001), powyższe zestawienie przyjmuje postać zaprezentowaną w Tabeli 1.

W związku z tym, rozwiązanie problemu wrażliwości miary może zostać oparte na dwóch alternatywnych strategiach. Pierwsza dotyczy wykazania, że określona miara jest niewrażliwa na przywołane typy problemów. Druga natomiast polega na podaniu dodatkowych warunków pozwalających wyróżnić wybrany typ miary potwierdzania. Drugi wariant zrealizowany został przez Carnapa. Podane przez niego warunki bazują na symetryczności miary potwierdzania. W kontekście niniejszego tekstu interesująca jest relacja pomiędzy postulatami dotyczącymi symetryczności a problemem (p4), który jako jedyny z listy dotyka miary potwierdzania c.

Tabela 1. Zestawienie miar potwierdzania z paradoksami BCT

	Czy miara jest odporna na określony problem?				
Problem / paradoks BCT	**d**	**r**	**l**	**c**	**s**
(p1) „Paradoks przechodniości potwierdzenia” (analiza J. Earman)	T	N	T	T	T
(p2) „Paradoks ziembieskości” (analiza E. Eells)	T	N	N	T	T
(p3) „Paradoks ziembieskości” (analiza E. Sober)	T	N	T	T	T
(p4) „Paradoks kruków” (analiza P. Horwich)	T	T	T	N	N

Źródło: Fitelson 2001: 24

3. Carnapowska miara potwierdzania c

Klasyfikacja probabilistycznego systemu Carnapa w obrębie współczesnej BCT związana jest z kontrowersjami natury terminologicznej. Powodem takiego stanu rzeczy jest częste nieprecyzyjne odnoszenie się do Bayesjańskiej interpretacji prawdopodobieństwa:

> Należy jasno zaznaczyć, że Carnap nie może być uznawany za zwolennika teorii, którą określamy współcześnie jako Bayesowska Interpretacja Prawdopodobieństwa, która traktuje prawdopodobieństwo w terminach racjonalnych stopni przekonań; może natomiast być traktowany jako pionier tzw. Bayesowskiej Teorii Potwierdzania, która utożsamia stopnie uzasadnienia z racjonalnymi stopniami przekonań (opartymi na określonych wytycznych) (Williamson 2010).

W *Logical Foundations of Probability* (1962) Carnap pozostawał na stanowisku, że istnieje jedna wyróżniona miara potwierdzania c. Aby ograniczyć zakres możliwych konstrukcji miar, nałożył dodatkowy warunek, bazujący na założeniu o symetrii. W rezultacie Carnapowska miara potwierdzania spełnia następujące typy symetrii:

a) symetryczność zdarzeń (**ES**):
$c\ (H, E \mid K) = -\ c\ (H, -E \mid K)$

b) symetryczność przemienna (**CS**):
$c\ (H, E \mid K) = c\ (E, H \mid K)$

c) symetryczność hipotez (**HS**):
$c\ (H, E \mid K) = -\ c\ (-H, E \mid K)$

d) symetryczność całkowita (**TS**):
$c\ (H, E \mid K) = c\ (-H, -E \mid K)$

Powyższe symetrie obrazują istotne zależności. Według (ES) zdarzenie E potwierdza hipotezę H w tym samym stopniu, co negacja E osłabia (*disconfirm*) hipotezę H. W świetle (HS) zdarzenie E potwierdza hipotezę H w ten sam sposób, jak to samo zdarzenie osłabia negację H. (TS) natomiast wynika z koniunkcji (ES) oraz (HS).

Jak więc widzimy, różne typy symetrii odpowiadają różnym relacjom pomiędzy hipotezą H a zdarzeniem E. Istotny jest fakt, że carnapowska miara potwierdzania c jako jedyna ze wszystkich nadmienionych (m1) – (m5) spełnia wszystkie typy symetryczności. Zdaniem Carnapa, warunki symetryczności pozwalają zobrazować prymarność miary c nad konkurencyjnymi rozwiązaniami, dlatego też określał ją mianem najlepszego kandydata na miarę potwierdzania.

Zdaniem wielu komentatorów symetryczności a) – d) są zbyt silnym warunkiem, a co za tym idzie, mimo iż pozwalają na uniknięcie pewnych typów paradoksów, zwiększają podatność na inne.

Przejrzyste zobrazowanie nakreślonej sytuacji prezentują Eells i Fitelson (2002), którzy przychylają się ku opinii o zbyt silnej symetrii. Ich zdaniem miara potwierdzania zastosowana w obrębie BTC powinna odrzucać symetrię typu (ES), (CS) oraz (TS). Uzasadnieniem takiego stanowiska jest poniższy przykład.

Załóżmy, że posiadamy bardzo wysoko uzasadnione przekonanie, że wszystkie kruki są czarne H*. Dodatkowo przyjmijmy, że hipoteza ta jest potwierdzana (bardzo słabo) przez kolejne przypadki nieczarnych niekruków oraz czarne niekruki. Ponadto wiemy, że informacja o nowo odkrytym obiekcie (nieznanym wcześniej ze względu na kolor), że nie jest nieczarnym krukiem, potwierdzi H*. Informację tę oznaczmy jako E*. Ponieważ hipoteza H* jest wysoko potwierdzona, zdarzenie E*, co nie powinno być zaskoczeniem, zapewnia potwierdzenie H*. ¬E* (informacja o tym, że obiekt jest nieczarnym krukiem) wydaje się bardzo mocno (ostatecznie) zaprzeczać H*. Tak więc ¬E*, jako bardzo silny dowód przeciwko H*, jest mocniejszy niż E* względem H*. W rezultacie uzyskujemy kontrprzykład względem symetryczności określonej mianem (ES).

Podobny, intuicyjny kontrprzykład można skonstruować wobec symetrii typu (CS). Wylosujmy kartę ze standardowej talii. Jako E oznaczmy zdarzenie, że wylosowana karta to siedem pik. Niech H będzie hipotezą, że karta jest czarna. Jest jasne, że E jest mocnym dowodem na korzyść H. Z drugiej strony ¬E (wylosowaną kartą nie jest siódemka pik) jest znikomo informatywne w stosunku H. Symetryczność relacji nie jest zachowana. W rezultacie okazuje się, że jedynym pożądanym typem symetrii jest (HS).

4. Konkluzja

Zaproponowany przez Carnapa warunek symetryczności okazuje się zbyt silny. Mimo iż tylko miara potwierdzania c jako jedyna spełnia wszystkie typy symetrii (ES), (HS), (CS), (TS), wyłącznie symetria typu (HS) nie podpada pod kontrprzykłady podobne do tych, jakie zaprezentowane zostały w poprzedniej części tekstu. Interesujący jest fakt, że pożądany typ symetryczności (HS) spełniony jest przez miarę d, dodatkowo miara ta nie spełnia symetryczności (ES), (CS), (TS), co mogłoby sugerować, że jest najlepszym kandydatem na efektywną miarę potwierdzania. Jednakże tu z kolei mamy do czynienia z innego typu obostrzeniami (nałożonymi na d), dyskredytującymi wskazaną miarę.

Bibliografia

Carnap R. (1962), *Logical Foundations of Probability*, 2nd ed., University of Chicago Press, Chicago.

Christensen D. (1999), *Measuring confirmation*, "Journal of Philosophy" XCVI, s. 437–461.

Earman J. (1992), *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*, Cambridge: MIT Press.

Eells E. (1982), *Rational Decision and Causality*, Cambridge: Cambridge University Press.

Eells E., Fitelson B. (2002), *Symmetries and asymmetries in evidential support*, "Philosophical Studies", vol. 107, no. 2, s. 129–142.

Fitelson B. (1999), *The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity*, "Philosophy of Science", vol. 66, s. 362–378.

Fitelson B. (2001), *Studies in Bayesian Confirmation Theory*, Ph.D. thesis, University of Wisconsin, Madison.

Good I. (1983), *Good Thinking: The Foundations of Probability and Its Applications*, Minneapolis: University of Minnesota Press.

Horwich P. (1982), *Probability and Evidence*, Cambridge: Cambridge University Press.

Joyce J. (1999), *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press.

Rosenkrantz R. (1994), *Bayesian confirmation: Paradise regained*, "The British Journal for the Philosophy of Science" 45, s. 467–476.

Sober E. (1994), *The principle of the common cause*, w: *From A Biological Point of View: Essays in Evolutionary Philosophy*, Cambridge: Cambridge University Press, s. 158–174,

Williamson J. (2010), *An Objective Bayesian Account of Confirmation*, University of Kent.

Streszczenie

Jednym z wyzwań Bayesjańskiej Teorii Potwierdzania jest rozwiązanie problemu wrażliwości miary. Dotyczy on relacji pomiędzy wyborem odpowiedniej miary potwierdzania oraz podatnością teorii na określone typy paradoksów. Jedną ze strategii przezwyciężenia powyższej sytuacji jest nałożenie dodatkowych warunków na miarę potwierdzania. Autorem jednej z propozycji jest Rudolf Carnap. Zaproponowane przez niego warunki bazują na symetryczności miary potwierdzania. Celem niniejszego tekstu jest analiza carnapowskich warunków symetryczności w kontekście problemu wrażliwości miary.