*Logika w zastosowaniach kognitywistycznych*

# Wprowadzenie do logiki epistemicznej. Przekonania i wiedza

(notatki do wykładów)

Andrzej Wiśniewski
Andrzej.Wisniewski@amu.edu.pl

*wersja beta 1.1*

(na podstawie: Wolfgang Lenzen: *Epistemic Logic*, w: Ilkka Niiniluoto, Matti Sintonen, Jan Woleński (eds.), **Handbook of Epistemology**, Kluwer Academic Publishers, Dordrecht/ Boston/ London 2004).

Rozważamy trzy modalności epistemiczne:

1) *a* ***jest całkowicie przekonany/ przekonana, że*** *p* [„is convinced that"]:

$$\mathbf{C}(a, p)$$

2) *a* ***jest przekonany/ przekonana, że*** *p* ['"believes that"]:

$$\mathbf{B}(a, p)$$

3) *a* ***wie, że*** *p* [„knows that"]:

$$\mathbf{K}(a, p)$$

## „Is convinced that”
## Przekonanie charakteryzujące się pewnością/ całkowite przekonanie

Symbol **C** będzie dalej rozumiany jako odnoszący się do przekonań charakteryzujących się pewnością i w tym właśnie sensie <u>całkowitych</u>; napis **C**(*a*, *p*) czytamy:

- „[podmiot] *a* jest całkowicie przekonany, że *p*”,

lub

- „[podmiot] *a* uważa, że z pewnością (zachodzi) *p*”

Całkowite przekonanie, **C**, uważa się zwykle za jedno z tzw. <u>nastawień sądzenionych</u> (ang. *propositional attitudes*). Pod pojęciem sądu (ang. *proposition*) rozumiemy tu zobiektywizowane znaczenie zdania oznajmującego.

Podobnie ujmuje się wiedzę, przekonanie nie charakteryzujące się pewnością, epistemiczne dopuszczanie, etc.

Używając pojęcia prawdopodobieństwa subiektywnego, ***Prob***, całkowite przekonanie możemy określić tak:

(**PROB-C**) $\mathbf{C}(a, p) \leftrightarrow \boldsymbol{Prob}(a, p) = 1.$

Prawdopodobieństwo – nawet subiektywne – dotyczy zdarzeń, natomiast wyżej stosowaliśmy pojęcie prawdopodobieństwa subiektywnego do sądów (w sensie ang. *propositions*). Aby miało to sens, należy założyć, iż:

(*) każdemu zdarzeniu odpowiada sąd (*proposition*), prezentujący/ opisujący to zdarzenie.

W logice epistemicznej[1] często używa się semantyki światów możliwych. Gdy weźmiemy model Kripkego <***W***, ***R***, ***V***> (dla normalnych *MRZ*), można przyjąć, że:

(**POSS-C**) $\boldsymbol{V}(\mathrm{C}(a, p), \boldsymbol{w}) = \mathbf{1}$ *wtw* $\forall \boldsymbol{w}^*(\boldsymbol{wRw}^* \rightarrow \boldsymbol{V}(p, \boldsymbol{w}^*) = \mathbf{1})$

gdzie **1** oznacza tym razem wartość logiczną, **Prawdę**. ***R*** będzie tutaj epistemiczną relacją alternatywności – o czym później.

[1] W szerokim sensie tego terminu. Czasami pod pojęciem logiki epistemicznej rozumie się tylko logikę wiedzy, natomiast logikę przekonań określa się mianem logiki doksastycznej (ang. *doxastic logic*).

Zarówno (**PROB-$\mathbf{C}$**), jak i (**POSS-$\mathbf{C}$**) dają w konsekwencji m.in. nastepujące prawa/zależności:

(**$\mathbf{C_1}$**) $\mathbf{C}(a, p) \wedge \mathbf{C}(a, q) \rightarrow \mathbf{C}(a, p \wedge q)$

(**$\mathbf{C_2}$**) $\mathbf{C}(a, p) \rightarrow \neg\mathbf{C}(a, \neg p)$

a także:

$$\neg(\mathbf{C}(a, p) \wedge \mathbf{C}(a, \neg p))$$

Niech $\mathbf{P}(a, p)$ oznacza:

- „[podmiot] *a* ***myśli, że jest możliwe to, że*** *p*"
  [„thinks *p* to be possible"].

Przyjmujemy następującą definicję:

**(Def. P)** $\mathbf{P}(a, p) \leftrightarrow \neg\mathbf{C}(a, \neg p)$

Oczywiście mamy:

$$\mathbf{C}(a, p) \rightarrow \mathbf{P}(a, p)$$

Korzystając z pojęcia prawdopodobieństwa subiektywnego ***Prob*** i z semantyki Kripkego, mielibyśmy/ mamy:

(**PROB-P**) $\mathbf{P}(a, p) \leftrightarrow \boldsymbol{Prob}(a, p) > 0.$

(**POSS-P**) $\boldsymbol{V}(\mathbf{P}(a, p), \boldsymbol{w}) = \mathbf{1}$ *wtw* $\exists \boldsymbol{w}^*(\boldsymbol{wRw}^* \wedge \boldsymbol{V}(p, \boldsymbol{w}^*) = \mathbf{1})$.

<u>Nie</u> zachodzi:

$$\mathbf{P}(a, p) \wedge \mathbf{P}(a, q) \to \mathbf{P}(a, p \wedge q)$$

ale zachodzi:

**($C_3$)** $$\mathbf{P}(a, p \wedge q) \to \mathbf{P}(a, p) \wedge \mathbf{P}(a, q)$$

Ponadto zachodzi:

**($C_4$)** $$\mathbf{C}(a, p \wedge q) \to \mathbf{C}(a, p) \wedge \mathbf{C}(a, q)$$

Ponieważ mieliśmy:

(**$C_1$**) $\mathbf{C}(a, p) \wedge \mathbf{C}(a, q) \to \mathbf{C}(a, p \wedge q)$

zatem dla $\mathbf{C}$ (ale nie dla $\mathbf{P}$!) mamy:

$$\mathbf{C}(a, p \wedge q) \leftrightarrow \mathbf{C}(a, p) \wedge \mathbf{C}(a, q)$$

Dla alternatywy dostajemy:

**($C_5$)** $\mathbf{C}(a, p) \vee \mathbf{C}(a, q) \rightarrow \mathbf{C}(a, p \vee q)$

**($C_6$)** $\mathbf{P}(a, p) \vee \mathbf{P}(a, q) \rightarrow \mathbf{P}(a, p \vee q)$

Zauważmy, że implikacja odwrotna do **($C_5$)** zachodzić nie powinna - i faktycznie nie zachodzi.

Sądy (*propositions*) to zobiektywizowane znaczenia zdań oznajmujących. Jeśli tak, to można uznać, że logicznie równoważne zdania wyrażają ten sam sąd i w konsekwencji prezentują /opisują to samo zdarzenie.

W efekcie zasadne staje się przyjęcie, iż w logice całkowitych przekonań obowiązuje następująca <u>reguła</u>:

**($C_7$)**

$$\frac{A \leftrightarrow B}{\mathbf{C}(a, A) \leftrightarrow \mathbf{C}(a, B)}$$

**Uwaga**: Z prawomocnym zastosowaniem powyższej reguły - a także innych omawianych na tym wykładzie reguł - mamy do czynienia tylko wówczas, gdy przesłanka jest <u>prawem</u> logiki, klasycznej lub epistemicznej. „Zwykła" prawdziwość przesłanki tutaj nie wystarcza!

Zwykle przyjmuje się również, że obowiązują reguły:

**($C_8$)**
$$\frac{A \rightarrow B}{\mathbf{C}(a, A) \rightarrow \mathbf{C}(a, B)}$$

**($C_9$)**
$$\frac{A}{\mathbf{C}(a, A)}$$

[Nawiasem mówiąc, rodzi to problemy, o których później.]

Przyjmujemy też, że obowiązuje *reguła podstawiania* oraz że możemy stosować reguły inferencyjne - pierwotne i wtórne - Klasycznego Rachunku Zdań. Dopuszczalne jest korzystanie z przesłanek będących podstawieniami tautologii KRZ.

Komentarz: Zapraszam na wykład :)

Dość intuicyjne (z uwagi na zasadę „uprzywilejowanego dostępu”/ *privileged access* do własnych stanów psychicznych) są następujące zależności:

**(E$_1$)** $\mathbf{C}(a, p) \rightarrow \mathbf{K}(a, \mathbf{C}(a, p))$

**(E$_2$)** $\neg\mathbf{C}(a, p) \rightarrow \mathbf{K}(a, \neg\mathbf{C}(a, p))$

**(E$_3$)** $\mathbf{K}(a, p) \rightarrow \mathbf{C}(a, p)$

skąd dostajemy prawa „iteracji” dla $\mathbf{C}$:

**(C$_{10}$)** $\mathbf{C}(a, p) \rightarrow \mathbf{C}(a, \mathbf{C}(a, p))$

**(C$_{11}$)** $\neg\mathbf{C}(a, p) \rightarrow \mathbf{C}(a, \neg\mathbf{C}(a, p))$

Gdy przyjmiemy, że **(C$_{10}$)** i **(C$_{11}$)** mogą być wzmocnione do równoważności, to każda iterowana $\mathbf{C}$-modalność jest równoważna albo formule postaci $\mathbf{C}(a, A)$, albo formule postaci $\neg\mathbf{C}(a, A)$, gdzie $A$ nie zawiera już operatora epistemicznego $\mathbf{C}$.

Natomiast zależność:

$$\mathbf{C}(a, p) \rightarrow p$$

<u>nie</u> zachodzi. Jest to intuicyjne: przekonanie, nawet połączone z (subiektywną) pewnością, że zachodzi *p,* nie gwarantuje, że zdarzenie opisywane przez *p* jest faktem. *Vide* posłowie RP lub też profesorowie IP :))

## „(Weakly) believes that”
## Przekonanie (niekoniecznie charakteryzujące się pewnością)

Symbol $\mathbf{B}$ odnosi się do przekonania (ang. *belief*), które nie musi - chociaż może! - łączyć się z pewnością. Napis $\mathbf{B}(a, p)$ będę krótko czytał:

- „[podmiot] *a* j***est przekonany, że*** *p*"

W ogólnym przypadku przekonanie nie musi być połączone z pewnością; wystarczy, że subiektywne prawdopodobieństwo tego, że *p*, jest większe od subiektywnego prawdopodobieństwa tego, że $\neg p$.

**(PROB-B)** $\mathbf{B}(a, p) \leftrightarrow \boldsymbol{Prob}(a, p) > ½$

Mamy, podobnie jak poprzednio:

**($B_1$)** $\mathbf{B}(a, p) \rightarrow \neg\mathbf{B}(a, \neg p)$

$\neg(\mathbf{B}(a, p) \wedge \mathbf{B}(a, \neg p))$.

Nie zachodzi:

$$\mathbf{B}(a, p) \wedge \mathbf{B}(a, q) \to \mathbf{B}(a, p \wedge q)$$

natomiast zachodzi:

**($E_4$)** $$\mathbf{B}(a, p) \wedge \mathbf{C}(a, q) \to \mathbf{B}(a, p \wedge q)$$

Mamy:

**($E_5$)** $$\mathbf{C}(a, p) \to \mathbf{B}(a, p)$$

ale nie mamy, i mieć nie powinniśmy, implikacji odwrotnej.

*Nota bene*, formuła $\mathbf{B}(a, p) \wedge \neg\mathbf{C}(a, p)$ wyraża to, że podmiot $a$ jest przekonany, że $p$, ale nie jest to przekonanie połączone z pewnością.

Ponadto mamy (z powodów podobnych, jak poprzednio):

**($E_6$)** $\mathbf{B}(a, p) \to \mathbf{K}(a, \mathbf{B}(a, p))$

**($E_7$)** $\neg\mathbf{B}(a, p) \to \mathbf{K}(a, \neg\mathbf{B}(a, p))$

co na mocy

**($E_3$)** $\mathbf{K}(a, p) \to \mathbf{C}(a, p)$

**($E_5$)** $\mathbf{C}(a, p) \to \mathbf{B}(a, p)$

daje prawa iteracji dla $\mathbf{B}$:

**($B_2$)** $\mathbf{B}(a, p) \to \mathbf{B}(a, \mathbf{B}(a, p))$

**($B_3$)** $\neg\mathbf{B}(a, p) \to \mathbf{B}(a, \neg\mathbf{B}(a, p))$

Podobnie jak dla $\mathbf{C}$ – i z analogicznych powodów – dla $\mathbf{B}$ obowiązuje reguła:

**($B_4$)**
$$\frac{A \leftrightarrow B}{\mathbf{B}(a, A) \leftrightarrow \mathbf{B}(a, B)}$$

Zazwyczaj przyjmuje się również, że obowiązują reguły:

**($B_5$)**
$$\frac{A \rightarrow B}{\mathbf{B}(a, A) \rightarrow \mathbf{B}(a, B)}$$

**($B_6$)**
$$\frac{A}{\mathbf{B}(a, A)}$$

oraz że dopuszczalne są przekształcenia, o których mówiliśmy w przypadku operatora $\mathbf{C}$ (zob. slajd 12).

Nie zachodzi:

$$\mathbf{B}(a, p) \rightarrow p$$

i tak powinno być :)

# „Knows that”
# Wiedza, że

Symbol $\mathbf{K}$ odnosi się do <u>wiedzy</u>. Napis $\mathbf{K}(a, p)$ czytamy:

- „[podmiot] *a* ***wie, że*** *p*”

Jakkolwiek rozumiemy $\mathbf{K}$, zachodzić musi:

**($K_1$)** $\mathbf{K}(a, p) \to p$

jako że wiedza, w przeciwieństwie do przekonania, pociąga prawdziwość/ faktyczność (ang. *factivity*).

Jak wiadomo, na pojęcie „wiedzy, że” można nakładać różne warunki – stąd też mamy wiele „logik wiedzy”. Gdy „wiedzę” rozumiemy jako „prawdziwe przekonanie charakteryzujące się pewnością” – a więc po platońsku – możemy wprowadzić operator wiedzy $\mathbf{K}^*$ definiowany tak:

**(Def.$\mathbf{K}^*$)** $\mathbf{K}^*(a, p) \leftrightarrow \mathbf{C}(a, p) \wedge p$

Oczywiście zachodzi:

**($\mathbf{K}^*_1$)** $\mathbf{K}^*(a, p) \rightarrow p$

a na mocy własności $\mathbf{C}$ dostajemy:

**($\mathbf{K}^*_2$)** $\mathbf{K}^*(a, p) \wedge \mathbf{K}^*(a, q) \rightarrow \mathbf{K}^*(a, p \wedge q)$

oraz to, że dla $\mathbf{K}^*$ obowiązują następujące reguły:

**(K*$_3$)** $$\frac{A \leftrightarrow B}{\mathbf{K}^*(a, A) \leftrightarrow \mathbf{K}^*(a, B)}$$

**(K*$_4$)** $$\frac{A \rightarrow B}{\mathbf{K}^*(a, A) \rightarrow \mathbf{K}^*(a, B)}$$

**(K*$_5$)** $$\frac{A}{\mathbf{K}^*(a, A)}$$

Zachodzi również:

**($K^*_6$)** $\mathbf{K}^*(a, p) \to \mathbf{K}^*(a, \mathbf{K}^*(a, p))$

Istotnie, ponieważ mamy:

**(Def.$\mathbf{K}^*$)** $\mathbf{K}^*(a, p) \leftrightarrow \mathbf{C}(a, p) \wedge p$

**($C_{10}$)** $\mathbf{C}(a, p) \to \mathbf{C}(a, \mathbf{C}(a, p))$

zatem jest tak, że:

$$\mathbf{K}^*(a, p) \to \mathbf{C}(a, \mathbf{C}(a, p)) \wedge \mathbf{C}(a, p) \wedge p$$

Z uwagi na:

$$\mathbf{C}(a, p \wedge q) \leftrightarrow \mathbf{C}(a, p) \wedge \mathbf{C}(a, q)$$

wnosimy stąd (odpowiednio podstawiając), iż:

$$\mathbf{K}^*(a, p) \to \mathbf{C}(a, \mathbf{C}(a, p) \wedge p) \wedge \mathbf{C}(a, p) \wedge p$$

Korzystając z definicji $\mathbf{K}^*$, dostajemy najpierw:

$$\mathbf{K}^*(a, p) \to \mathbf{K}^*(a, \mathbf{C}(a, p) \wedge p)$$

a stąd:

$$\mathbf{K}^*(a, p) \to \mathbf{K}^*(a, \mathbf{K}^*(a, p))$$

Natomiast zależność:

**($K^*_6$)** $\neg\mathbf{K}^*(a, p) \rightarrow \mathbf{K}^*(a, \neg\mathbf{K}^*(a, p))$

<u>nie</u> zachodzi dla $\mathbf{K}^*$. Tak zresztą powinno być: gdy *p* nie jest prawdą, natomiast podmiot epistemiczny *a* jest przekonany, że *p* jest prawdą, to jest tak, że $\neg\mathbf{K}^*(a, p)$, ale nie musi być tak, że $\mathbf{K}^*(a, \neg\mathbf{K}^*(a, p))$.

[Jak zobaczymy dalej, z określonych powodów powyższa zależność jest jednak prawem pewnych logik wiedzy.]

Dla $\mathbf{K}^*$ mamy jednak:

**($E_8$)** $\neg\mathbf{C}(a, p) \rightarrow \mathbf{K}^*(a, \neg\mathbf{C}(a, p))$

Analizę operatora epistemicznego „...wie, że...” można wzbogacać w różnych kierunkach. Można ją też – jak zobaczymy (ale nie na tym wykładzie) osłabiać.

Pamiętajmy jednak, ze gdy mówimy o wiedzy, zachodzić musi:

**($K_1$)** $\mathbf{K}(a, p) \rightarrow p$

Zwykle przyjmuje się, że obowiązują reguły:

**($K_3$)** $$\frac{A \leftrightarrow B}{\mathbf{K}(a, A) \leftrightarrow \mathbf{K}(a, B)}$$

**($K_4$)** $$\frac{A \rightarrow B}{\mathbf{K}(a, A) \rightarrow \mathbf{K}(a, B)}$$

**($K_5$)**

$$\frac{A}{\mathbf{K}(a, A)}$$

oraz że obowiązuje reguła podstawiania, a także inne - pierwotne i wtórne - reguły inferencyjne Klasycznego Rachunku Zdań.

Często przyjmuje się, że zachodzi:

**($K_6$)** $\mathbf{K}(a, p) \to \mathbf{K}(a, \mathbf{K}(a, p))$

Można dodatkowo z góry żądać, aby zachodziła następująca zależność:

**($K_2$)** $\mathbf{K}(a, p) \wedge \mathbf{K}(a, q) \to \mathbf{K}(a, p \wedge q)$

(w przypadku $\mathbf{K}^*$ odpowiedniki **($K_2$)** i **($K_6$)** były następstwami definicji operatora $\mathbf{K}^*$; teraz tego nie zakładamy), oraz zależność:

$$\mathbf{K}(a, p) \to \neg\mathbf{K}(a, \neg p).$$

Związki między $\mathbf{C}$, $\mathbf{B}$ i $\mathbf{K}$ można widzieć tak:

**($E_9$)** $\mathbf{C}(a, p) \to \mathbf{B}(a, \mathbf{K}(a, p))$

**($E_{10}$)** $\mathbf{C}(a, p) \to \mathbf{C}(a, \mathbf{K}(a, p))$

Gdyby wzmocnić **($E_{10}$)** do równoważności, tj. gdyby przyjąć:

**($E_{10}$)’** $\mathbf{C}(a, p) \leftrightarrow \mathbf{C}(a, \mathbf{K}(a, p))$

to w oparciu o wzmocnione do równoważności **($C_{10}$)**, tj.:

**($C_{10}$)’** $\mathbf{C}(a, p) \leftrightarrow \mathbf{C}(a, \mathbf{C}(a, p))$

można udowodnić:

**($E_{11}$)** $\mathbf{C}(a, \mathbf{C}(a, p)) \leftrightarrow \mathbf{C}(a, \mathbf{K}(a, p))$

**($E_{11}$)** orzeka, iż charakteryzujące się pewnością przekonanie oraz wiedza są <u>subiektywnie nieodróżnialne</u>.

Jak już wspomniałem, jest rzeczą dyskusyjną, czy powinniśmy przyjmować:

**($K_6$)** $\neg\mathbf{K}(a, p) \rightarrow \mathbf{K}(a, \neg\mathbf{K}(a, p))$

Czasami proponuje się przyjęcie zależności (nieco) słabszych:

**($K_7$)** $\neg\mathbf{K}(a, \neg\mathbf{K}(a, p)) \rightarrow \mathbf{K}(a, \neg\mathbf{K}(a, \neg\mathbf{K}(a, p)))$

**($K_8$)** $p \wedge \neg\mathbf{K}(a, \neg\mathbf{K}(a, p)) \rightarrow \mathbf{K}(a, p)$

Sens intuicyjny powyższych formuł nieco rozjaśnia następująca zależność:

**($E_{12}$)** $\neg\mathbf{K}(a, \neg\mathbf{K}(a, p)) \leftrightarrow \mathbf{C}(a, p)$

którą udowodnimy sobie na ćwiczeniach :)

W charakterze dygresji warto dodać, że odpowiedniki **($K_7$)** i **($K_8$)** dla $\mathbf{K}^*$:

**($K^*_7$)** $\neg\mathbf{K}^*(a, \neg\mathbf{K}^*(a, p)) \rightarrow \mathbf{K}^*(a, \neg\mathbf{K}^*(a, \neg\mathbf{K}^*(a, p)))$

**($K^*_8$)** $p \wedge \neg\mathbf{K}^*(a, \neg\mathbf{K}^*(a, p)) \rightarrow \mathbf{K}^*(a, p)$

można udowodnić.

Gdyby chcieć charakteryzować $\mathbf{K}$ w kategoriach semantyki Kripkego, mielibyśmy:

(**POSS-K**) $\boldsymbol{V}(\mathbf{K}(a, p), \boldsymbol{w}) = \mathbf{1}$ *wtw* $\forall \boldsymbol{w}^*(\boldsymbol{wRw}^* \rightarrow \boldsymbol{V}(p, \boldsymbol{w}^*) = \mathbf{1})$

czyli warunek przypominający odpowiedni warunek dla $\mathbf{C}$. Różnica przejawia się w warunkach nakładanych na relację epistemicznej alternatywności $\boldsymbol{R}$. W szczególności, aby otrzymać:

$$\mathbf{K}p \rightarrow p$$

powinniśmy przyjąć, że $\boldsymbol{R}$ jest zwrotna – czego z kolei nie wolno nam przyjmować w semantykach dla $\mathbf{C}$ czy $\mathbf{B}$. O tym jednak – i o epistemicznych rachunkach zdań – na następnych wykładach.