

Logika w zastosowaniach kognitywistycznych

Wprowadzenie do logiki epistemicznej.

Przekonania i wiedza

(notatki do wykładów)

Andrzej Wiśniewski
Andrzej.Wisniewski@amu.edu.pl

wersja beta 1.0

(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”]:

$C(a, p)$

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

$B(a, p)$

3) *a wie, że p* [„knows that”]:

$K(a, p)$

„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 całkowitych; 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. nastawień sądzeniowych (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:

$$\text{(PROB-C)} \quad C(a, p) \leftrightarrow \text{Prob}(a, p) = 1.$$

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 dokładnie jeden sąd (*proposition*), prezentujący/ opisujący to zdarzenie.

W logice epistemicznej¹ często używa się semantyki światów możliwych. Gdy weźmiemy model Kripkego $\langle W, R, V \rangle$ (dla normalnych *MRZ*), można przyjąć, że:

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

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 doksyastycznej (ang. *doxastic logic*).

Zarówno (**PROB-C**), jak i (**POSS-C**) dają w konsekwencji m.in. następujące prawa/zależności:

$$(C_1) \quad C(a, p) \wedge C(a, q) \rightarrow C(a, p \wedge q)$$

$$(C_2) \quad C(a, p) \rightarrow \neg C(a, \neg p)$$

a także:

$$\neg(C(a, p) \wedge C(a, \neg p))$$

Niech $P(a, p)$ oznacza:

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

Przyjmujemy następującą definicję:

$$\text{(Def. P)} \quad P(a, p) \leftrightarrow \neg C(a, \neg p)$$

Oczywiście mamy:

$$C(a, p) \rightarrow P(a, p)$$

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

$$\text{(PROB-P)} \quad P(a, p) \leftrightarrow \text{Prob}(a, p) > 0.$$

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

Nie zachodzi:

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

ale zachodzi:

$$(\mathbf{C}_3) \quad \mathbf{P}(a, p \wedge q) \rightarrow \mathbf{P}(a, p) \wedge \mathbf{P}(a, q)$$

Ponadto zachodzi:

$$(\mathbf{C}_4) \quad \mathbf{C}(a, p \wedge q) \rightarrow \mathbf{C}(a, p) \wedge \mathbf{C}(a, q)$$

Ponieważ mieliśmy:

$$(\mathbf{C}_1) \quad \mathbf{C}(a, p) \wedge \mathbf{C}(a, q) \rightarrow \mathbf{C}(a, p \wedge q)$$

zatem dla **C** (ale nie dla **P**!) mamy:

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

Dla alternatywy dostajemy:

$$(C_5) \quad C(a, p) \vee C(a, q) \rightarrow C(a, p \vee q)$$

$$(C_6) \quad P(a, p) \vee P(a, q) \rightarrow P(a, p \vee q)$$

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 przekonania obowiązuje następująca reguła:

$$(C_7) \quad \frac{A \leftrightarrow B}{C(a, A) \leftrightarrow C(a, B)}$$

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

Zwykle przyjmuje się również, że obowiązują reguły:

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

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

[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:

$$(E_1) \quad C(a, p) \rightarrow K(a, C(a, p))$$

$$(E_2) \quad \neg C(a, p) \rightarrow K(a, \neg C(a, p))$$

$$(E_3) \quad K(a, p) \rightarrow C(a, p)$$

skąd dostajemy prawa „iteracji” dla C :

$$(C_{10}) \quad C(a, p) \rightarrow C(a, C(a, p))$$

$$(C_{11}) \quad \neg C(a, p) \rightarrow C(a, \neg C(a, p))$$

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

Natomiast zależność:

$$C(a, p) \rightarrow p$$

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

- „[podmiot] *a* **jest 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 $\neg p$.

$$\text{(PROB-B)} \quad \mathbf{B}(a, p) \leftrightarrow \text{Prob}(a, p) > \frac{1}{2}$$

Mamy, podobnie jak poprzednio:

$$\begin{aligned} \text{(B}_1\text{)} \quad & \mathbf{B}(a, p) \rightarrow \neg \mathbf{B}(a, \neg p) \\ & \neg(\mathbf{B}(a, p) \wedge \mathbf{B}(a, \neg p)). \end{aligned}$$

Nie zachodzi:

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

natomiast zachodzi:

$$(E_4) \quad \mathbf{B}(a, p) \wedge \mathbf{C}(a, q) \rightarrow \mathbf{B}(a, p \wedge q)$$

Mamy:

$$(E_5) \quad \mathbf{C}(a, p) \rightarrow \mathbf{B}(a, p)$$

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):

$$(E_6) \quad \mathbf{B}(a, p) \rightarrow \mathbf{K}(a, \mathbf{B}(a, p))$$

$$(E_7) \quad \neg \mathbf{B}(a, p) \rightarrow \mathbf{K}(a, \neg \mathbf{B}(a, p))$$

co na mocy

$$(E_3) \quad \mathbf{K}(a, p) \rightarrow \mathbf{C}(a, p)$$

$$(E_5) \quad \mathbf{C}(a, p) \rightarrow \mathbf{B}(a, p)$$

daje prawa iteracji dla **B**:

$$(B_2) \quad \mathbf{B}(a, p) \rightarrow \mathbf{B}(a, \mathbf{B}(a, p))$$

$$(B_3) \quad \neg \mathbf{B}(a, p) \rightarrow \mathbf{B}(a, \neg \mathbf{B}(a, p))$$

Podobnie jak dla **C** – i z analogicznych powodów – dla **B** obowiązuje reguła:

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

Zazwyczaj przyjmuje się również, że obowiązują reguły:

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

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

oraz że dopuszczalne są przekształcenia, o których mówiliśmy w przypadku operatora **C** (zob. slajd 12).

Nie zachodzi:

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

i tak powinno być :)

**„Knows that”
Wiedza, że**

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

- „[podmiot] *a wie, że p*”

Jakkolwiek rozumiemy **K**, zachodzić musi:

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

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:

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

Oczywiście zachodzi:

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

a na mocy własności \mathbf{C} dostajemy:

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

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

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

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

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

Zachodzi również:

$$(K^*_6) \quad K^*(a, p) \rightarrow K^*(a, K^*(a, p))$$

Istotnie, ponieważ mamy:

$$(Def.K^*) \quad K^*(a, p) \leftrightarrow C(a, p) \wedge p$$

$$(C_{10}) \quad C(a, p) \rightarrow C(a, C(a, p))$$

zatem jest tak, że:

$$K^*(a, p) \rightarrow C(a, C(a, p)) \wedge C(a, p) \wedge p$$

Z uwagi na:

$$C(a, p \wedge q) \leftrightarrow C(a, p) \wedge C(a, q)$$

wnosimy stąd (odpowiednio podstawiając), iż:

$$K^*(a, p) \rightarrow C(a, C(a, p) \wedge p) \wedge C(a, p) \wedge p$$

Korzystając z definicji K^* , dostajemy najpierw:

$$K^*(a, p) \rightarrow K^*(a, C(a, p) \wedge p)$$

a stąd:

$$K^*(a, p) \rightarrow K^*(a, K^*(a, p))$$

Natomiast zależność:

$$(K^*_6) \quad \neg K^*(a, p) \rightarrow K^*(a, \neg K^*(a, p))$$

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

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

Dla K^* mamy jednak:

$$(E_8) \quad \neg C(a, p) \rightarrow K^*(a, \neg C(a, p))$$

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

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

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

Zwykle przyjmuje się, że obowiązują reguły:

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

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

$$(K_5) \quad \frac{A}{\mathbf{K}(a, A)}$$

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

Często przyjmuje się, że zachodzi:

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

Można dodatkowo z góry żądać, aby zachodziła następująca zależność:

$$(K_2) \quad \mathbf{K}(a, p) \wedge \mathbf{K}(a, q) \rightarrow \mathbf{K}(a, p \wedge q)$$

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

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

Związki między **C**, **B** i **K** można widzieć tak:

$$(E_9) \quad C(a, p) \rightarrow B(a, K(a, p))$$

$$(E_{10}) \quad C(a, p) \rightarrow C(a, K(a, p))$$

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

$$(E_{10}') \quad C(a, p) \leftrightarrow C(a, K(a, p))$$

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

$$(C_{10}') \quad C(a, p) \leftrightarrow C(a, C(a, p))$$

można udowodnić:

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

(E_{11}) orzeka, iż charakteryzujące się pewnością przekonanie oraz wiedza są subiektywnie nieodróżnialne.

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

$$(K_6) \quad \neg K(a, p) \rightarrow K(a, \neg K(a, p))$$

Czasami proponuje się przyjęcie zależności (nieco) słabszych:

$$(K_7) \quad \neg K(a, \neg K(a, p)) \rightarrow K(a, \neg K(a, \neg K(a, p)))$$

$$(K_8) \quad p \wedge \neg K(a, \neg K(a, p)) \rightarrow K(a, p)$$

Sens intuicyjny powyższych formuł nieco rozjaśnia następująca zależność:

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

którą udowodnimy sobie na ćwiczeniach :)

W charakterze dygresji warto dodać, że odpowiedniki (K_7) i (K_8) dla K^* :

$$(K^*_7) \quad \neg K^*(a, \neg K^*(a, p)) \rightarrow K^*(a, \neg K^*(a, \neg K^*(a, p)))$$

$$(K^*_8) \quad p \wedge \neg K^*(a, \neg K^*(a, p)) \rightarrow K^*(a, p)$$

można udowodnić.

Gdyby chcieć charakteryzować **K** w kategoriach semantyki Kripkego, mielibyśmy:

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

czyli warunek przypominający odpowiedni warunek dla **C**. Różnica przejawia się w warunkach nakładanych na relację epistemicznej alternatywności **R**. W szczególności, aby otrzymać:

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

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