

Logika w zastosowaniach kognitywistycznych

**Wprowadzenie do logiki epistemicznej.
Zdaniowa logika przekonań KDE4
i jej semantyka.
Część II**

(notatki do wykładów)

Andrzej Wiśniewski
Andrzej.Wisniewski@amu.edu.pl

wersja beta 1.6

**Semantyka Meyera - van der Hoeka dla
KDE4 i KE4**

W książce J. J. Ch. Meyera i W. van der Hoek'a pt. *Epistemic Logic for AI and Computer Science* (Kluwer, Dordrecht 1995) znajdujemy bardzo interesujący wariant semantyki światów możliwych dla epistemicznych rachunków przekonań **KDE4/KD45** i **KE4/K45**.

Rozpocznijmy od przypadku rachunku **KDE4**.

Rozważmy strukturę postaci:

(#) $\langle W, \{w_0\}, R \rangle$

gdzie **W** jest zbiorem („światów możliwych”), **w₀** to, intuicyjnie rzecz biorąc, „nasz świat”, **w₀ ∉ W**, a **R** jest binarną relacją w iloczynie kartezjańskim:

$$W \cup \{w_0\} \times W \cup \{w_0\}.$$

Model określony na strukturze $\langle W, \{w_0\}, R \rangle$ ma postać:

(##) $\langle W, \{w_0\}, R, V \rangle$

gdzie V jest funkcją wartościowania, przyporządkowującą każdej parze uporządkowanej, której pierwszym elementem jest formuła, a drugim element zbioru $W \cup \{w_0\}$, dokładnie jedną z wartości logicznych $1, 0$; funkcję tę definiujemy nieomalże w standardowy sposób, podobnie jak pojęcia prawdziwości formuły w modelu postaci (##) i prawdziwości formuły w strukturze postaci (#).

„Nieomalże”, albowiem warunki dla formuł postaci PA oraz CA są następujące:

- $V(PA, w) = 1$ wtw istnieje $w^* \in W \cup \{w_0\}$ takie, że wRw^* oraz $V(A, w^*) = 1$.
- $V(CA, w) = 1$ wtw dla każdego $w^* \in W \cup \{w_0\}$ takiego, że wRw^* : $V(A, w^*) = 1$.

Prawdziwość w modelu postaci (##) to prawdziwość w każdym świecie ze zbioru $W \cup \{w_0\}$. Prawdziwość w strukturze postaci (#) to prawdziwość w każdym modelu (rozważanego typu) opartym na tej strukturze.

Rozważmy teraz model $M = \langle W, \{w_0\}, R, V \rangle$ analizowanego rodzaju spełniający następujące warunki:

(a₁) $w_0 R w$ dla każdego $w \in W$ („każdy świat z W jest epistemicznie alternatywny względem ‘naszego’ świata w_0 ”),

(a₂) $w R w^*$ dla dowolnych $w, w^* \in W$ („wszystkie światy epistemicznie alternatywne względem ‘naszego’ świata w_0 są epistemicznie alternatywne względem siebie”).

Uwaga: Należy pamiętać, że na mocy definicji $w_0 \notin W$.

Zauważmy, że gdy $\langle w_0, w_0 \rangle \notin R$, to wartość logiczna formuły postaci CA w „naszym” świecie w_0 modelu M nie zależy od wartości logicznej formuły A w świecie w_0 tego modelu. Zależy ona natomiast od wartości logicznych, jakie przyjmuje formuła A w każdym ze światów ze zbioru W modelu M . Proste rozumowanie pokazuje, że dla (każdego) modelu M analizowanego rodzaju spełniającego powyższe warunki zachodzi:

o jeśli $\langle w_0, w_0 \rangle \notin R$, to $V(CA, w_0) = 1$ wtw dla każdego $w \in W$:

$$V(A, w) = 1.$$

Można udowodnić, iż:

*Formuła A jest tezą rachunku przekonań **KDE4** wtedy i tylko wtedy, gdy formuła A jest prawdziwa w każdej strukturze postaci $\langle \mathbf{W}, \{w_0\}, \mathbf{R} \rangle$, w której \mathbf{W} jest zbiorem niepustym, a relacja \mathbf{R} spełnia łącznie następujące warunki:*

(a₁) $w_0 \mathbf{R} w$ dla każdego $w \in \mathbf{W}$ („każdy świat z \mathbf{W} jest epistemicznie alternatywny względem ‘naszego’ świata w_0 ”),

(a₂) $w \mathbf{R} w^$ dla dowolnych $w, w^* \in \mathbf{W}$ („wszystkie światy epistemicznie alternatywne względem ‘naszego’ świata w_0 są epistemicznie alternatywne względem siebie”).*

Uwaga: Należy pamiętać, że na mocy poprzednich ustaleń $w_0 \notin \mathbf{W}$, natomiast \mathbf{R} jest binarną relacją w iloczynie kartezjańskim:

$$\mathbf{W} \cup \{w_0\} \times \mathbf{W} \cup \{w_0\}.$$

W przypadku rachunku **KE4** sytuacja jest prawie taka sama; różnica polega na tym, że nie wymagamy, aby zbiór **W** każdej wyjściowej struktury był niepusty. Mamy zatem:

*Formuła A jest tezą rachunku przekonań **KE4** wtedy i tylko wtedy, gdy formuła A jest prawdziwa w każdej strukturze postaci $\langle W, \{w_0\}, R \rangle$, w której relacja **R** spełnia łącznie następujące warunki:*

- (a₁) $w_0 R w$ dla każdego $w \in W$ („każdy świat z **W** jest epistemicznie alternatywny względem ‘naszego’ świata w_0 ”),*
- (a₂) $w R w^*$ dla dowolnych $w, w^* \in W$ („wszystkie światy epistemicznie alternatywne względem ‘naszego’ świata w_0 są epistemicznie alternatywne względem siebie”).*

Komentarze: zapraszam na wykład :)

Przykłady zastosowań

Przypuśćmy teraz, że a jest podmiotem epistemicznym (*epistemic agent*), posiadającym m.in. przekonania żywione z całkowitą pewnością. Formuła postaci CA będzie zatem znaczyć tyle, co:

- **podmiot a jest całkowicie przekonany** (*convinced*), że A ,

podobnie formuła postaci PA znaczy, że:

- **podmiot a myśli, że jest możliwe, że A** (*thinks A to be possible*).

Założmy, że operator epistemiczny C jest w rozważanym przypadku rozumiany zgodnie z logiką **KDE4** (tj. „**KDE4** jest logiką całkowitych przekonań podmiotu a ”).

Niech w_0 oznacza świat rzeczywisty – „nasz świat”.

Rozważmy trzy zdania:

- (1) *Jan jest logikiem.*
- (2) *Jan jest psychastenikiem.*
- (3) *Jan jest paranoikiem.*

Przypuśćmy, że dla podmiotu *a* światami epistemicznie alternatywnymi wobec w_0 („możliwymi scenariuszami wydarzeń”) są - wzajemnie epistemicznie alternatywne - światy w_1 oraz w_2 takie, że:

<i>świat w_1</i>	<i>świat w_2</i>
zachodzi (1)	zachodzi (1)
nie zachodzi (2)	zachodzi (2)
zachodzi (3)	nie zachodzi (3)

Powstaje pytanie: **o jakich zdaniach możemy prawomocnie stwierdzić, że wyrażają one całkowite przekonania podmiotu *a*?** Pytanie to możemy sprowadzić do szeregu pytań postaci:

(*) *Czy zdanie ‘CA’ jest prawdą w świecie w_0 ?*

albowiem interesują nas (całkowite) przekonania podmiotu *a* w „naszym” świecie w_0 .

Przyjmując, że zdania (1) – (3) są kolejno reprezentowane przez zmienne zdaniowe p , q , r , możemy powiedzieć, że powyższa sytuacja jest reprezentowana przez – **dowolny** - model (w sensie Meyera i van der Hoeka) $M^{\$}$:

$$\langle \{w_1, w_2\}, \{w_0\}, R, V \rangle$$

spełniający następujące warunki:

$$(a) R = \{ \langle w_0, w_1 \rangle, \langle w_0, w_2 \rangle, \langle w_1, w_2 \rangle, \langle w_2, w_1 \rangle, \langle w_1, w_1 \rangle, \langle w_2, w_2 \rangle \}$$

$$(b) V(p, w_1) = 1; V(q, w_1) = 0; V(r, w_1) = 1;$$

$$(c) V(p, w_2) = 1; V(q, w_2) = 1; V(r, w_2) = 0.$$

Odpowiednikiem pytania o schemacie (*) jest teraz pytanie mające schemat:

$$(**) \quad \text{Czy } V(CA, w_0) = 1 ?$$

jako że – przypomnijmy – formuła jest prawdziwa w świecie modelu wtw wartość tej formuły (w tym świecie tego modelu) wynosi 1.

Niech $R \rightarrow w_i$ oznacza zbiór wszystkich światów modelu M^s alternatywnych względem świata w_i tego modelu.

I. Zapytajmy:

Czy $V(Cp, w_0) = 1$?

Odpowiedź: **tak**, bo $R \rightarrow w_0 = \{w_1, w_2\}$, a $V(p, w_1) = 1$ oraz $V(p, w_2) = 1$.

II. Zapytajmy:

Czy $V(Cq, w_0) = 1$?

Odpowiedź: **nie**, bo $R \rightarrow w_0 = \{w_1, w_2\}$, a $V(q, w_1) = 0$.

III. Zapytajmy:

Czy $V(Cr, w_0) = 1$?

Odpowiedź: **nie**, bo $R \rightarrow w_0 = \{w_1, w_2\}$, a $V(r, w_2) = 0$.

IV. Zapytajmy:

Czy $V(Pq, w_0) = 1$?

Odpowiedź: **tak**, bo $R \rightarrow w_0 = \{w_1, w_2\}$, a $V(q, w_2) = 1$.

V. Zapytajmy:

Czy $V(\mathbf{P}r, \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo $R \rightarrow \mathbf{w}_0 = \{\mathbf{w}_1, \mathbf{w}_2\}$, a $V(r, \mathbf{w}_1) = 1$.

VI. Zapytajmy:

Czy $V(\mathbf{C}(q \vee r), \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo $R \rightarrow \mathbf{w}_0 = \{\mathbf{w}_1, \mathbf{w}_2\}$, a skoro $V(r, \mathbf{w}_1) = 1$, to również $V(q \vee r, \mathbf{w}_1) = 1$; podobnie skoro $V(q, \mathbf{w}_2) = 1$, to również $V(q \vee r, \mathbf{w}_2) = 1$.

VII. Zapytajmy:

Czy $V(\mathbf{C}(p \rightarrow q \vee r), \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo $R \rightarrow \mathbf{w}_0 = \{\mathbf{w}_1, \mathbf{w}_2\}$, a mamy $V(q \vee r, \mathbf{w}_1) = 1$ oraz $V(q \vee r, \mathbf{w}_2) = 1$.

VIII. Zapytajmy:

Czy $V(\mathbf{C}q \vee \mathbf{C}r, \mathbf{w}_0) = 1$?

Odpowiedź: **nie**, bo $V(\mathbf{C}q, \mathbf{w}_0) = 0$ oraz $V(\mathbf{C}r, \mathbf{w}_0) = 0$.

Uwaga: Zagadnienia (**VIII** i **IX**) mają odmienny charakter od poprzednich. Dlaczego? Zapraszam na wykład :)

IX. Zapytajmy:

Czy $V(\neg \mathbf{C}q \rightarrow \mathbf{C}r, \mathbf{w}_0) = 1$?

Odpowiedź: **nie**, bo $V(\mathbf{C}q \vee \mathbf{C}r, \mathbf{w}_0) = 0$.

Przy okazji: przykłady **VI** i **VIII** pokazują, że nie jest tak, iż tezą rachunku **KDE4** jest dowolna formuła o postaci $\mathbf{C}(A \vee B) \rightarrow \mathbf{C}A \vee \mathbf{C}B$!

X. Zapytajmy:

Czy $V(\mathbf{CC}p, \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo tezą rachunku **KDE4** jest każda formuła postaci:

$$\mathbf{CA} \rightarrow \mathbf{CCA}$$

a więc także formuła:

$$\mathbf{C}p \rightarrow \mathbf{CC}p.$$

Jeśli tak, to formuła $\mathbf{C}p \rightarrow \mathbf{CC}p$ jest prawdziwa w rozważanym modelu $\mathbf{M}^\$,$ skąd wnosimy, że jest ona prawdziwa w każdym świecie tego modelu – a więc także w „naszym” świecie \mathbf{w}_0 . Skoro jednak mieliśmy (por. przykład I) $V(\mathbf{C}p, \mathbf{w}_0) = 1$, to mamy również $V(\mathbf{CC}p, \mathbf{w}_0) = 1$.

XI. Zapytajmy:

Czy $V(P_{\neg}q, w_0) = 1$?

Odpowiedź: **tak**, bo $V(Cq, w_0) = 0$, czyli $V(\neg Cq, w_0) = 1$. Tezą rachunku **KDE4** jest każda formuła postaci:

$$\neg CA \rightarrow P_{\neg}A$$

czyli także formuła:

$$\neg Cq \rightarrow P_{\neg}q.$$

Tak więc skoro mamy $V(\neg Cq, w_0) = 1$, to mamy też $V(P_{\neg}q, w_0) = 1$.

XII. Zapytajmy:

Czy $V(P_{\neg}r, w_0) = 1$?

Odpowiedź: **tak**, bo $V(Cr, w_0) = 0$, a dalej rozumujemy analogicznie jak w poprzednim przypadku.

XIII. Zapytajmy:

Czy $V(C\neg Cq, \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo mamy $V(Cq, \mathbf{w}_0) = 0$, czyli $V(\neg Cq, \mathbf{w}_0) = 1$, a tezę rachunku **KDE4** jest każda formuła postaci:

$$\neg CA \rightarrow C\neg CA.$$

XIV. Zapytajmy:

Czy $V(C\neg Cr, \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo mamy $V(Cr, \mathbf{w}_0) = 0$.

XV. Zapytajmy:

Czy $V(C(\neg(p \wedge \neg p))), \mathbf{w}_0) = 1$?

Odpowiedź: **tak**, bo w rachunku **KDE4** obowiązuje **RG**.

XVI. Zapytajmy:

Czy $V(\mathbf{C}((p \vee q) \wedge (p \vee r)), \mathbf{w}_0) = \mathbf{1}$?

Odpowiedź: **tak**, bo tezą rachunku **KDE4** jest każda formuła postaci:

$$(!) \quad \mathbf{C}A \wedge \mathbf{C}(B \vee C) \rightarrow \mathbf{C}((A \vee B) \wedge (A \vee C))$$

czyli także formuła:

$$(!*) \quad \mathbf{C}p \wedge \mathbf{C}(q \vee r) \rightarrow \mathbf{C}((p \vee q) \wedge (p \vee r))$$

a mamy już $V(\mathbf{C}p, \mathbf{w}_0) = \mathbf{1}$ (por. przykład **I**) oraz $V(\mathbf{C}(q \vee r), \mathbf{w}_0) = \mathbf{1}$ (zob. przykład **VI**).

Gdybyśmy jednak nie wiedzieli/nie pamiętali, że formuły postaci (!) są tezami rachunku **KDE4**, odpowiedzi na zadane pytanie możemy szukać budując tabelę analityczną dla formuły (!*) lub – co na jedno wychodzi - sprawdzając metodą tabel analitycznych, czy formuła:

$$\mathbf{C}((p \vee q) \wedge (p \vee r))$$

wynika na gruncie rachunku **KDE4** ze zbioru formuł:

$$\{\mathbf{C}p, \mathbf{C}(q \vee r)\}.$$

Tabela 1.

$$\neg(\mathbf{C}p \wedge \mathbf{C}(q \vee r) \rightarrow \mathbf{C}((p \vee q) \wedge (p \vee r))), 0$$

$$\mathbf{C}p \wedge \mathbf{C}(q \vee r), 0$$

$$\neg\mathbf{C}((p \vee q) \wedge (p \vee r)), 0$$

$$\mathbf{C}p, 0$$

$$\mathbf{C}(q \vee r), 0$$

$$\mathbf{P}\neg((p \vee q) \wedge (p \vee r)), 0$$

OR1

$$\neg((p \vee q) \wedge (p \vee r)), 1$$

$$\neg(p \vee q), 1$$

$$\neg p, 1$$

$$\neg q, 1$$

$$p, 1$$

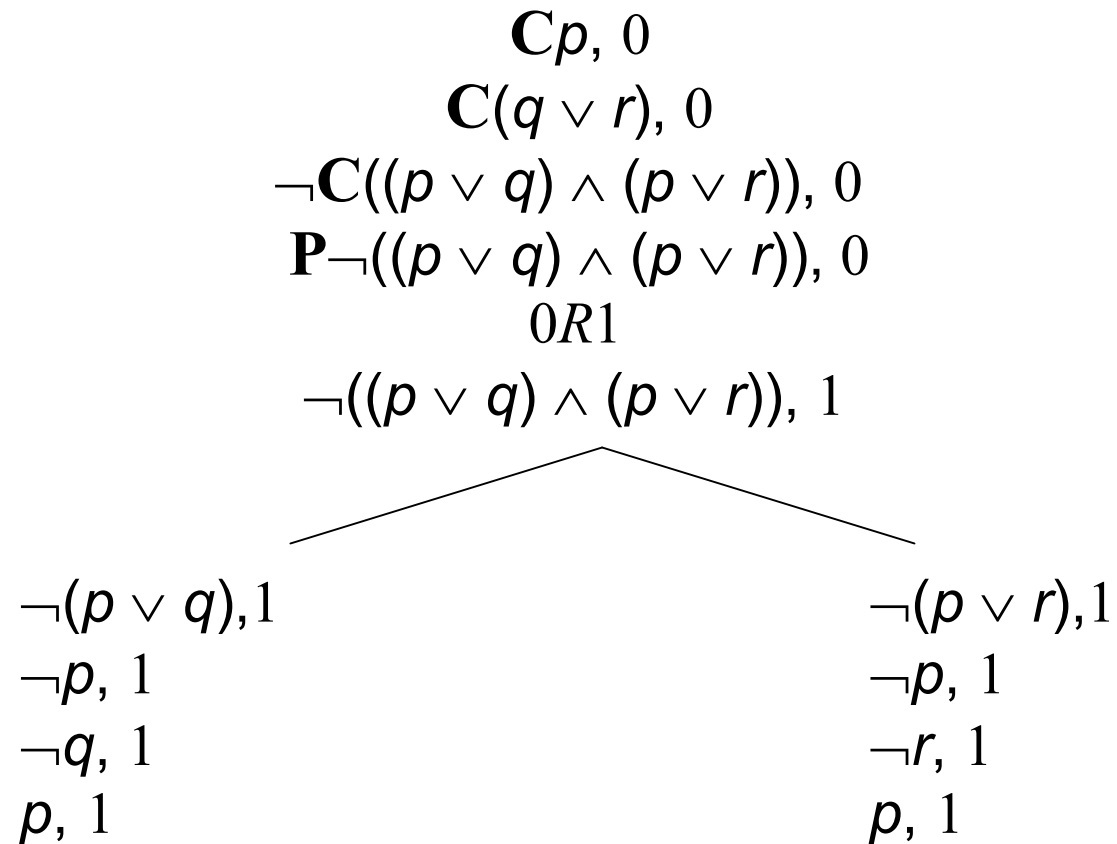
$$\neg(p \vee r), 1$$

$$\neg p, 1$$

$$\neg r, 1$$

$$p, 1$$

Tabela 2.



Podsumujmy, czego udało nam się jak dotąd dowiedzieć o podmiocie epistemicznym a :

a jest całkowicie przekonany, że:	nie jest tak, że a jest całkowicie przekonany, że:	a myśli, że jest możliwe, że:
p	q	q
$q \vee r$	r	r
$p \rightarrow q \vee r$		$\neg q$
$(p \vee q) \wedge (p \vee r)$		$\neg r$
$\neg(p \wedge \neg p)$		
$\mathbf{C}p$		
$\neg\mathbf{C}q$		
$\neg\mathbf{C}r$		

Przypomnijmy, że p : ‘Jan jest logikiem’, q : ‘Jan jest psychastenikiem”,
a r : ‘Jan jest paranoikiem’.

Nasza analiza doprowadziła do scharakteryzowanie pewnych nastawień sądzeniowych (*propositional attitudes*) podmiotu *a*. Ściślej rzecz biorąc, **przypisaliśmy** podmiotowi *a* te nastawienia, *zakładając*, że **KDE4** jest „właściwą” logiką całkowitych przekonań.

Przypisując podmiotowi *a* nastawienia sądzeniowe wobec zdań/ sądów (1) – (3) oraz pewnych zdań/sądów z nich zbudowanych, opieraliśmy się na informacji wyjściowej dotyczącej tego, jaką postać mają „alternatywy epistemiczne” rozważane/ dopuszczane przez podmiot *a*. Informacja ta *nie dotyczyła* natomiast tego, jakie są nastawienia sądzeniowe podmiotu *a* wobec zdań/ sądów (1) – (3) oraz odpowiednich zdań/sądów z nich utworzonych.

Zauważmy na koniec, że logikę **KDE4** - podobnie jak inne logiki epistemiczne – można również stosować na inne sposoby. Przedstawię tutaj jeden z nich.

Przypuśćmy, że mamy informację odnośnie tego, że rozważany podmiot a jest całkowicie przekonany co do (sądów) A_1, \dots, A_n . **Pytania, które sobie stawiamy, dotyczą tego, czy podmiot a jest również całkowicie przekonany, że B_1, \dots, B_k i/ lub czy podmiot myśli, że jest możliwe, że D_1, \dots, D_m .** Szukając odpowiedzi, zakładamy, że **KDE4** jest „właściwą logiką przekonań” i stawiamy pytania typu:

*Czy formuła CB_i wynika na gruncie logiki **KDE4** ze zbioru formuł $\{CA_1, \dots, CA_n\}$?*

*Czy formuła PD_i wynika na gruncie logiki **KDE4** ze zbioru formuł $\{CA_1, \dots, CA_n\}$?*

gdzie $1 \leq i \leq k, m$; rozważamy rzecz jasna formuły języka rachunku **KDE4** reprezentujące odpowiednie sądy.

Odpowiedzi poszukujemy stosując np. metodę tabel analitycznych dla **KDE4**.

Semantyczne uzasadnienie przyjętego sposobu postępowania jest następujące: wynikanie na gruncie rachunku **KDE4** gwarantuje „transmisję prawdziwości” formuł względem dowolnego świata dowolnego modelu – a zatem również z uwagi na świat w_0 reprezentujący „nasz świat”.

Zilustrujmy rzecz przykładem.

Przypuśćmy, że podmiot a jest całkowicie przekonany, że:

p : ‘Szef jest idiotą’.

$p \rightarrow q$: ‘Jeśli szef jest idiotą, to stanowią dla niego zagrożenie’.

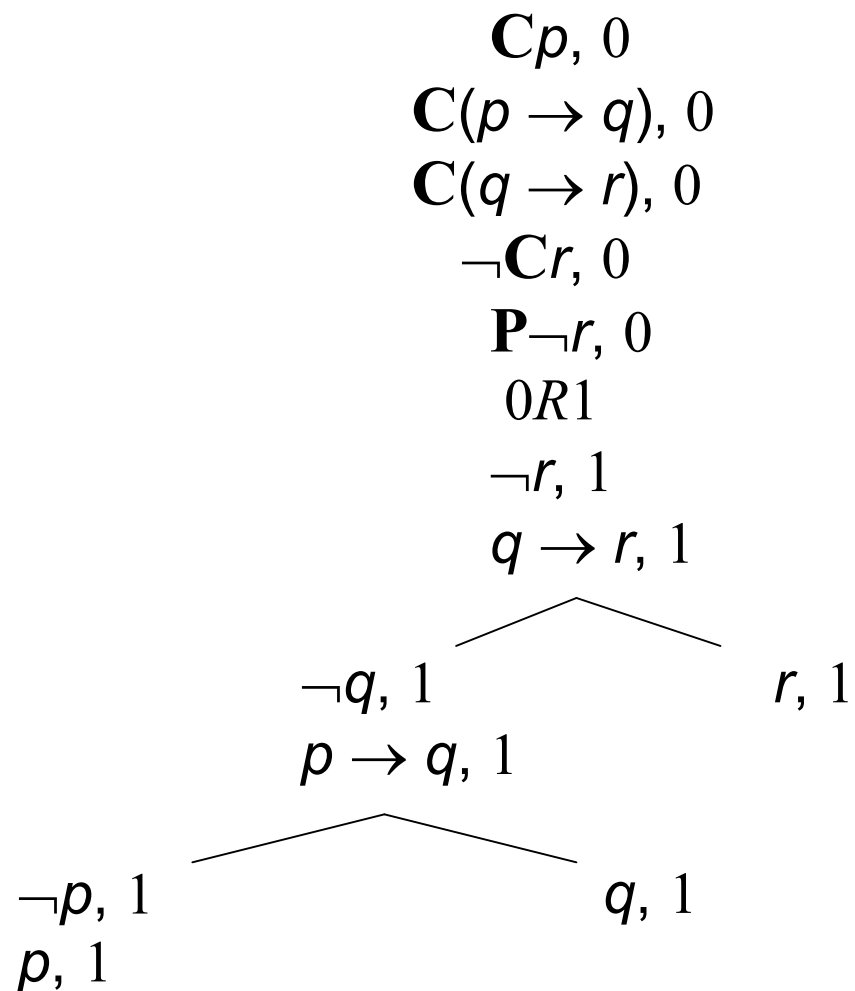
$q \rightarrow r$: ‘Jeśli stanowią zagrożenie dla szefa, to szef prędzej czy później zwolni mnie z pracy’.

Czy w tej sytuacji podmiot a jest również całkowicie przekonany, że szef prędzej czy później zwolni go z pracy ?

Zatem:

Czy jest tak, że $\{Cp, C(p \rightarrow q), C(q \rightarrow r)\} \vdash_{\text{KDE4}} Cr$?

Budujemy tabelę analityczną¹:



¹ Chyba że pamiętamy, iż $\mathbf{C}p \wedge (\mathbf{C}(p \rightarrow q) \wedge \mathbf{C}(q \rightarrow r)) \rightarrow \mathbf{C}r$ jest tezą logiki **KDE4**.

Tabela jest zamknięta, tak więc **możemy przypisać** podmiotowi *a* całkowite przekonanie, że szef prędzej czy później zwolni go/ ją z pracy.

Czy podmiot *a* faktycznie żywi takie przekonanie, czy też raczej jest zobowiązany, aby je żywić (*is committed to believe*) – to sprawa odrębna.

Więcej na ten temat – na wykładzie.