

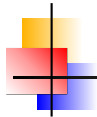


Logică epistemică

Cursul opțional Elemente de securitate și logică aplicată,
Anul 3, Semestrul 2, 2022

Laurențiu Leuștean

Web page: <http://cs.unibuc.ro/~lleustean/>



LOGICI PENTRU CUNOAȘTERE



Logici pentru cunoaștere

- ▶ au fost dezvoltate pentru a raționa despre **sisteme multiagent**.
- ▶ sunt folosite pentru a demonstra proprietăți ale acestor sisteme.
- ▶ sunt folosite pentru a reprezenta și raționa despre **informația** pe care agenții o posedă: **cunoașterea** lor.



- ▶ Întrebarea **Ce este un agent?** nu are un răspuns definitiv.
- ▶ S-au dat multe răspunsuri, unele dintre ele contradictorii, inconsistente.

Definiția din **Michael Wooldridge, An Introduction to MultiAgent Systems, Second Edition**, John Wiley & Sons, 2009:

Un **agent** este un sistem informatic capabil de **a acțiune independentă (autonomă)** în numele utilizatorului sau proprietarului său. Un agent realizează singur ceea ce trebuie făcut pentru a îndeplini obiectivele pentru care a fost proiectat, mai degrabă decât să i se spună constant ce are de făcut.



Ce este un sistem multiagent?

Definiția din [Michael Wooldridge, An Introduction to MultiAgent Systems, Second Edition](#), John Wiley & Sons, 2009:

Un **sistem multiagent** constă dintr-un număr de agenți care **interacționează**. Agenții acționează în numele unor utilizatori cu scopuri și motivații diferite. Pentru a interacționa cu succes, ei trebuie să **coopereze**, să **se coordoneze** și să **negocieze** unul cu celălalt, la fel ca oamenii.

Definiția din [Yoav Shoham, Kevin Leyton-Brown, Multiagents Systems](#), Cambridge University Press, 2009:

Un **sistem multiagent** este un sistem care include multiple entități **autonome** care au informații divergente și/sau interese divergente.



Motivația pentru studiul sistemelor multiagent provine din interesul pentru **agenți artificiali (software sau hardware)**, de exemplu agenți software de pe Internet.

Subiectul este foarte **interdisciplinar**.

Exemple

- ▶ roboți autonomi;
- ▶ agenți comerciali;
- ▶ agenți care asistă (sau înlocuiesc) jucători umani într-un joc multiplayer;
- ▶ agenți de interfață - care facilitează interacțiunea între utilizator și diversele resurse computaționale;
- ▶ ...



Exemplu

- ▶ Considerăm un sistem distribuit, în care mai multe procesoare realizează în mod autonom unele calcule comune.
- ▶ Desigur, natura comună a calcululelor înseamnă că procesoarele trebuie să comunice între ele.
- ▶ Un set de probleme apare atunci când comunicarea este predispusă la erori. În acest caz, analistul de sistem poate spune ceva de genul următor:
 - ▶ Procesorul A a trimis mesajul către procesorul B . Este posibil ca mesajul să nu ajungă, iar procesorul A știe acest lucru.
 - ▶ Mai mult, procesorul A știe că procesorul B știe că A știe că, dacă a fost trimis un mesaj, este posibil să nu ajungă.
- ▶ Vrem să facem precise astfel de raționamente.

Adesea modelarea se face în contextul unor probleme stilizate, cu o poveste distractivă asociată.



Puzzle-ul cu copiii noroioși (Muddy children puzzle)

- ▶ Un grup de n copii intră în casă după ce s-au jucat afară în noroi. Sunt întâmpinați pe hol de tatăl lor, care observă că p dintre copii au noroi pe frunte.
- ▶ Tatăl face următorul anunț: “Cel puțin unul dintre voi are noroi pe frunte.”
- ▶ Fiecare copil poate vedea frunțile tuturor celorlalți copii, dar nu și pe a sa.
- ▶ Apoi tatăl spune: “Știe cineva dintre voi că are noroi pe frunte? Dacă știe, să ridice mâna.”
- ▶ Nimeni nu ridică mâna.
- ▶ Tatăl repetă întrebarea, din nou nimeni nu ridică mâna.
- ▶ Tatăl nu cedează și continuă să repete întrebarea.
- ▶ După exact p repetiții, toți copiii cu noroi pe frunte ridică mâna simultan.



Puzzle-ul cu copiii noroioși - $p = 1$

- ▶ Există un singur copil cu noroi pe frunte.
- ▶ Acesta știe că ceilalți copii sunt curați.
- ▶ Când tatăl zice că cel puțin un copil este noroi, copilul noroios concluzionează că trebuie să fie el.
- ▶ Niciunul dintre ceilalți copii nu știe în acest moment dacă este sau nu noroi.
- ▶ Copilul noroios ridică mâna după prima întrebare a tatălui.
- ▶ După ce acesta ridică mâna, ceilalți copii știu că sunt curați.



Puzzle-ul cu copiii noroioși - $p = 2$

- ▶ Există doi copii cu noroi pe frunte.
- ▶ Imaginează-ți că tu ești unul din cei doi copii noroioși.
- ▶ Vezi că unul din ceilalți copii este noroios.
- ▶ După primul anunț al tatălui, nu ai suficiente informații pentru a ști dacă ai noroi pe frunte. S-ar putea să fii noroios, dar se poate și ca celălalt copil să fie singurul cu noroi pe frunte.
- ▶ Prin urmare, nu ridici mâna după prima întrebare a tatălui.
- ▶ Observi că celălalt copil noroios nu ridică mâna.
- ▶ Îți dai seama că și tu trebuie să fii noroios, altfel acel copil ar fi ridicat mâna.
- ▶ Așadar, după a doua întrebare a tatălui, ridici mâna. Desigur, la fel face și celălalt copil noroios.



Puzzle-ul cu copiii noroioși

- ▶ Acest argument se poate extinde la $p = 3, 4, \dots$
- ▶ Desigur, ar fi de preferat o teoremă generală care se aplică tuturor p .
- ▶ Pentru aceasta avem nevoie de un **model formal de "cunoaștere"** care se aplică la acest exemplu.

Exemple din limbajul natural

Vrem să reprezentăm afirmații precum:

- ▶ "Ion știe că plouă."
- ▶ "Maria știe că Ion știe că va ploua mâine."
- ▶ "Atât Maria cât și Ion știu că plouă."
- ▶ "Andrei știe că Cronos este tatăl lui Zeus."



Considerăm următoarea afirmație

Andrei știe că Cronos este tatăl lui Zeus.

Cum o putem formaliza în logica de ordinul întâi?

O încercare este

$$\varphi := \textit{\textsf{Stie}}(\textit{\textsf{Andrei}}, \textit{\textsf{Tată}}(\textit{\textsf{Zeus}}, \textit{\textsf{Cronos}})),$$

unde *Știe* și *Tată* sunt simboluri de relații binare, iar *Andrei*, *Zeus*, *Cronos* sunt simboluri de constante.

O problemă sintactică

φ nu este formulă a logicii de ordinul întâi, deoarece *Tată*(*Zeus*, *Cronos*) este o formulă, nu un termen.



O problemă semantică

- ▶ Considerăm un alt simbol de constantă *Jupiter*, care denumește, ca și *Zeus*, deitatea supremă a lumii antice clasice.
- ▶ Este natural atunci să considerăm că e adevărată următoarea formulă:

$$\psi := (\text{Zeus} = \text{Jupiter}).$$

- ▶ Fie

$$\chi := \text{\textit{\textit{Știe}}}(\text{Andrei}, \text{Tată}(\text{Jupiter}, \text{Cronos})).$$

- ▶ Dacă φ , χ ar fi formule din logica de ordinul întâi, then, din adevărul lui ψ se poate infera că φ și χ sunt echivalente.
- ▶ Intuiția însă respinge această inferență: a ști că tatăl lui Zeus este Cronos **nu este totuna** cu a crede că tatăl lui Jupiter este Cronos.



O problemă semantică

Problema este că, în general, valoarea de adevăr a enunțului

Andrei știe p .

nu depinde doar de valoarea de adevăr a lui p . Cu alte cuvinte,
cunoașterea nu este adevăr-funcțională.

- ▶ Logica de ordinul întâi nu este potrivită în forma sa standard pentru a raționa despre cunoaștere.
- ▶ Sunt necesare formalisme alternative.

Domeniul **logicilor pentru cunoaștere** a fost inițiat cu publicarea, în 1962, a cărții **Knowledge and Belief. An Introduction to the Logic of the Two Notions** al cărei autor este lui **Jaakko Hintikka**.



Soluția problemei sintactice:

- ▶ Se folosesc **logici modale**, al căror limbaj conține **operatori modali**, care nu sunt adevăr-funcționali.
- ▶ Un exemplu de astfel de logici sunt **logicile epistemice**.

Soluția problemei semantice:

- ▶ Se folosește **semantica lumilor posibile**, propusă inițial de Hintikka pentru logica epistemică.
- ▶ Cunoașterea unui agent este caracterizată folosind o mulțime de **lumi posibile** (numite de Hintikka **alternative epistemice**), cu o relație de **accesibilitate** între ele.
- ▶ Ceva adevărat în **toate** alternativele epistemice ale agentului nostru e considerat ca fiind cunoscut de agent.



Cea mai comună formulare a semanticii lumilor posibile în logica modală (normală) a fost dezvoltată de Saul Kripke:

- ▶ Saul Kripke, *Semantical analysis of logica modală I. Normal modal propositional calculi*, Zeitschrift für Mathematische Logik și Grundlagen der Mathematik, 9 (1963), 67-96.



LOGICI MODALE



Definiția 2.1

Limbajul modal de bază ML_0 este format din:

- ▶ o mulțime $PROP$ de *propoziții atomice* sau *variabile propoziționale* (notate p, q, r, v, \dots);
- ▶ conectorii propoziționali: \neg, \rightarrow ;
- ▶ parantezele: $(,)$;
- ▶ operatorul modal \Box (se citește *cutie*).

Mulțimea $Sim(ML_0)$ a *simbolurilor* lui ML_0 este

$$Sim(ML_0) := PROP \cup \{\neg, \rightarrow, (,), \Box\}.$$

Expresiile lui ML_0 sunt șirurile finite de simboluri ale lui ML_0 .



Definiția 2.2

Formulele limbajului modal de bază ML_0 sunt expresiile definite astfel:

- (F0) Orice variabilă propozițională este formulă.
- (F1) Dacă φ este formulă, atunci $(\neg\varphi)$ este formulă.
- (F2) Dacă φ și ψ sunt formule, atunci $(\varphi \rightarrow \psi)$ este formulă.
- (F3) Dacă φ este formulă, atunci $(\Box\varphi)$ este formulă.
- (F4) Numai expresiile obținute aplicând regulile (F0), (F1), (F2), (F3) sunt formule.

Observația 2.3

Formulele lui ML_0 sunt definite, folosind notația Backus-Naur, astfel:

$$\varphi ::= p \mid (\neg\varphi) \mid (\varphi \rightarrow \psi) \mid (\Box\varphi), \quad \text{unde } p \in PROP.$$



Limbajul modal de bază

- ▶ În practică, renunțăm la parantezele exterioare, le punem numai atunci când sunt necesare. Scriem $\neg\varphi$, $\varphi \rightarrow \psi$, $\Box\varphi$.
- ▶ Operatorul modal \Box are precedență mai mare decât conectorii propoziționali.
- ▶ \neg are precedență mai mare decât \rightarrow .

Conectorii propoziționali \vee , \wedge , \leftrightarrow și constantele \top (adevăru**l**), \perp (falsu**l**) sunt definiți ca în logica propozițională clasică:

$$\varphi \vee \psi := \neg\varphi \rightarrow \psi$$

$$\varphi \wedge \psi := \neg(\varphi \rightarrow \neg\psi)$$

$$\varphi \leftrightarrow \psi := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi) \quad \top := p \rightarrow p, \quad \perp := \neg\top.$$

Operatorul modal dual

Dualul lui \Box se notează \Diamond (se citește **diamant**) și se definește astfel:

$$\Diamond\varphi := \neg\Box\neg\varphi.$$



Logica modală clasică

În logica modală clasică, $\Box\varphi$ este citit ca **este necesar φ** . Atunci $\Diamond\varphi$ înseamnă **nu este necesar ca non φ** , adică **este posibil ca φ** .

Exemple de formule pe care le putem privi ca principii corecte includ:

- ▶ $\Box\varphi \rightarrow \Diamond\varphi$ (**ce este necesar este și posibil**)
- ▶ $\varphi \rightarrow \Diamond\varphi$ (**ce este, este posibil**).

Ce putem spune despre formule ca $\varphi \rightarrow \Box\Diamond\varphi$ (**ce este, este necesar posibil**) sau $\Diamond\varphi \rightarrow \Box\Diamond\varphi$ (**ce este posibil, este necesar posibil**)? Pot fi ele privite ca adevăruri generale? Pentru a da un răspuns la astfel de întrebări trebuie să definim o semantică pentru logica modală clasică.



Logica epistemică

În logica epistemică, limbajul modal de bază este folosit pentru a raționa despre cunoaștere. În această logică,

$\Box\varphi$ se citește **agentul știe că φ** .

Se scrie $K\varphi$ în loc de $\Box\varphi$.

Deoarece discutăm despre cunoaștere, este natural să considerăm că este adevărată formula

$K\varphi \rightarrow \varphi$ (**dacă agentul știe că φ , atunci φ trebuie să aibă loc**)

Presupunând că agentul nu este atotștiutor, formula $\varphi \rightarrow K\varphi$ ar trebui să fie falsă.



Semantica



Definiția 2.4

O **structură relațională** este un tuplu \mathcal{F} format din:

- ▶ o mulțime nevidă W , numită **universul** (sau **domeniul**) lui \mathcal{F} ;
- ▶ o mulțime de relații pe W .

Presupunem că fiecare structură relațională conține cel puțin o relație. Elementele lui W se numesc **stări**, **lumi**, **puncte**, **noduri**, **timpi**, **instanțe** sau **situații**.

Exemplul 2.5

- ▶ $\mathcal{F} = (W, R)$, unde R este o relație de ordine parțială pe W (adică o relație binară pe W care este reflexivă, antisimetrică și tranzitivă).
- ▶ $\mathcal{F} = (W, R)$, unde R este o relație de echivalență pe W (adică o relație binară pe W care este reflexivă, simetrică și tranzitivă).



Sistemele de tranziții etichetate (*Labeled Transition Systems*), sau, mai simplu, **sistemele de tranziții**, sunt structuri relaționale simple, foarte folosite în informatică.

Definiția 2.6

*Un sistem de tranziții este o pereche $(W, \{R_a \mid a \in A\})$, unde W este o mulțime nevidă de **stări**, A este o mulțime nevidă de **etichete** și, pentru fiecare $a \in A$, $R_a \subseteq W \times W$ este o relație binară pe W .*

Sistemele de tranziții pot fi văzute ca modele abstracte de calcul: stările sunt stările posibile ale unui calculator, etichetele sunt programe și $(u, v) \in R_a$ înseamnă că există o execuție a programului a care începe în starea u și se termină în starea v .



Fie W o mulțime nevidă și $R \subseteq W \times W$ o relație binară.

Scriem de obicei Rwv în loc de $(w, v) \in R$. Dacă Rwv , atunci spunem că v este **R -accesibil** din w .

Inversa lui R se notează R^{-1} și se definește astfel:

$$R^{-1}vw \quad \text{ddacă} \quad Rvw.$$

Definim $R^n (n \geq 0)$ inductiv:

$$R^0 = \{(w, w) \mid w \in W\}, R^1 = R, R^{n+1} = R \circ R^n.$$

Așadar, pentru orice $n \geq 2$, avem că $R^n wv$ ddacă există u_1, \dots, u_{n-1} a.î $Rwu_1, Ru_1u_2, \dots, Ru_{n-1}v$.



În continuare dăm semantica limbajului modal de bază cu ajutorul structurilor relaționale. Facem acest lucru în două moduri:

- ▶ la nivelul **modelelor**, unde definim noțiunile fundamentale de **satisfacere** și **adevăr**;
- ▶ la nivelul **cadrelor**, care ne permite să definim noțiunea cheie de **validitate**.



Definiția 2.7

Un **cadru Kripke** (Kripke frame) pentru ML_0 este o pereche $\mathcal{F} = (W, R)$ astfel încât

- ▶ W este o mulțime nevidă;
- ▶ R este o relație binară pe W .

Așadar, un cadru Kripke este pur și simplu o structură relațională cu o singură relație binară. Elementele lui W se numesc **stări** sau **lumi**.

Interpretare folosind agenți

R_{wv} dacă agentul consideră lumea v posibilă, conform informațiilor pe care le are în lumea w . Gândim R ca o relație de **posibilitate**, deoarece R definește ce lumi sunt considerate posibile de către agent din orice lume dată.

Definiția 2.8

Un **model Kripke** (Kripke model) pentru ML_0 este o pereche $\mathcal{M} = (\mathcal{F}, V)$, unde

- ▶ $\mathcal{F} = (W, R)$ este un cadru pentru ML_0 ;
- ▶ $V : PROP \rightarrow 2^W$ este o funcție numită **evaluare**.

Prin urmare, funcția V asignează oricărei propoziții atomice $p \in PROP$ submulțimea $V(p)$ a lui W . Informal, ne gândim la $V(p)$ ca la mulțime stărilor în care p este adevărată.

Fie $\mathcal{F} = (W, R)$ un cadru Kripke și $\mathcal{M} = (\mathcal{F}, V)$ un model Kripke. Spunem că modelul $\mathcal{M} = (\mathcal{F}, V)$ este **bazat pe** cadrul $\mathcal{F} = (W, R)$ sau că \mathcal{F} este cadrul **suport** al lui \mathcal{M} . Dacă $w \in W$, scriem uneori $w \in \mathcal{F}$ sau $w \in \mathcal{M}$.

Scriem de cele mai multe ori $\mathcal{M} = (W, R, V)$.

Definim în continuare ce înseamnă că o formulă este adevărată într-o stare dintr-un model Kripke. Adevărul depinde atât de model cât și de stare.

Definiția 2.9

Fie $\mathcal{M} = (W, R, V)$ un model Kripke și w o stare în \mathcal{M} . Definim inductiv noțiunea

formula φ este **satisfăcută** (sau **adevărată**) în \mathcal{M} în starea w ,
notație $\mathcal{M}, w \Vdash \varphi$

- $\mathcal{M}, w \Vdash p$ *ddacă* $w \in V(p)$, unde $p \in PROP$
- $\mathcal{M}, w \Vdash \neg\varphi$ *ddacă* nu este adevărat că $\mathcal{M}, w \Vdash \varphi$
- $\mathcal{M}, w \Vdash \varphi \rightarrow \psi$ *ddacă* $\mathcal{M}, w \Vdash \varphi$ implică $\mathcal{M}, w \Vdash \psi$
- $\mathcal{M}, w \Vdash \Box\varphi$ *ddacă* pentru orice $v \in W$,
 Rwv implică $\mathcal{M}, v \Vdash \varphi$.



Notăție

Dacă \mathcal{M} nu satisface φ în w , scriem $\mathcal{M}, w \not\models \varphi$ și spunem că φ este **falsă** în \mathcal{M} în starea w .

Pentru orice stare $w \in W$,

- ▶ $\mathcal{M}, w \models \neg\varphi$ ddacă $\mathcal{M}, w \not\models \varphi$.
- ▶ $\mathcal{M}, w \models \varphi \rightarrow \psi$ ddacă $(\mathcal{M}, w \not\models \varphi \text{ sau } \mathcal{M}, w \models \psi)$.

Clauza pentru \Box are ca inspirație filosofia lui Leibniz:

- ▶ **necesitate** înseamnă **adevăr în toate lumile posibile**.
- ▶ **posibilitate** înseamnă **adevăr într-o lume posibilă**.

Fie $\mathcal{M} = (W, R, V)$ un model Kripke.

Propoziția 2.10

Pentru orice stare w în \mathcal{M} și orice formule φ, ψ ,

$$\mathcal{M}, w \not\models \perp$$

$$\mathcal{M}, w \models \top$$

$$\mathcal{M}, w \models \varphi \wedge \psi \quad \text{dacă} \quad \mathcal{M}, w \models \varphi \text{ și } \mathcal{M}, w \models \psi$$

$$\mathcal{M}, w \models \varphi \vee \psi \quad \text{dacă} \quad \mathcal{M}, w \models \varphi \text{ sau } \mathcal{M}, w \models \psi$$

$$\mathcal{M}, w \models \Diamond \varphi \quad \text{dacă} \quad \text{există } v \in W \text{ a.î. } R w v \text{ și } \mathcal{M}, v \models \varphi.$$

Dem.: Exercițiu.

Definiția 2.11

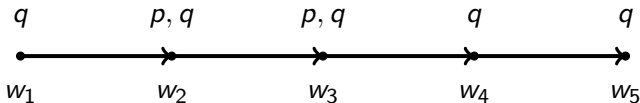
- ▶ O formulă φ este **global adevărată** sau **universal adevărată** în \mathcal{M} dacă $\mathcal{M}, w \models \varphi$ pentru orice $w \in W$. **Notăție:** $\mathcal{M} \models \varphi$
- ▶ O formulă φ este **satisfiabilă** în \mathcal{M} dacă există o stare $w \in W$ a.î. $\mathcal{M}, w \models \varphi$.



Un model Kripke $\mathcal{M} = (W, R, V)$ poate fi reprezentat ca un graf etichetat:

- ▶ nodurile grafului sunt stările modelului.
- ▶ etichetăm fiecare nod cu propozițiile atomice care sunt adevărate în acel nod.
- ▶ există un arc de la nodul w la nodul v ddacă Rwv .

Exemplu

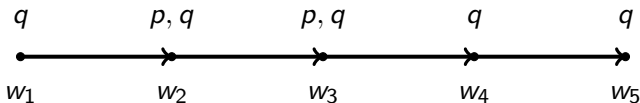


Avem că $\mathcal{M} = (W, R, V)$, unde $W = \{w_1, w_2, w_3, w_4, w_5\}$; Rw_iw_j ddacă $j = i + 1$; $\Phi = \{p, q, r\}$, $V(p) = \{w_2, w_3\}$, $V(q) = \{w_1, w_2, w_3, w_4, w_5\}$ și $V(r) = \emptyset$.



Exemplu

Fie $\mathcal{M} = (W, R, V)$ modelul Kripke reprezentat prin:



Demonstrați că

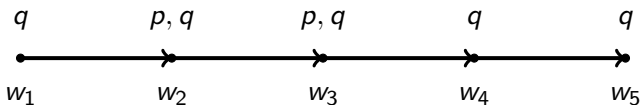
- (i) $\mathcal{M}, w_1 \Vdash \Diamond \Box p$.
- (ii) $\mathcal{M}, w_1 \nVdash \Diamond \Box p \rightarrow p$.
- (iii) $\mathcal{M}, w_2 \Vdash \Diamond(p \wedge \neg r)$.
- (iv) $\mathcal{M}, w_1 \Vdash q \wedge \Diamond(q \wedge \Diamond(q \wedge \Diamond(q \wedge \Diamond(q))))$.
- (v) $\mathcal{M} \Vdash \Box q$.

Dem.: (i) $\mathcal{M}, w_1 \Vdash \Diamond \Box p$ ddacă există $v \in W$ a.î. $Rw_1 v$ și $\mathcal{M}, v \Vdash \Box p$. Luăm $v := w_2$. Cum $Rw_1 w_2$, rămâne să demonstrăm că $\mathcal{M}, w_2 \Vdash \Box p$.



Exemplu

Fie $\mathcal{M} = (W, R, V)$ modelul Kripke reprezentat prin:



Dem.: (continuare) Avem că

$$\begin{aligned} \mathcal{M}, w_2 \Vdash \Box p &\iff \text{pentru orice } u \in W, R w_2 u \text{ implică } \mathcal{M}, u \Vdash p. \\ &\iff \mathcal{M}, w_3 \Vdash p \text{ (deoarece } w_3 \text{ este unicul } u \in W \\ &\quad \text{a.î. } R w_2 u) \\ &\iff w_3 \in V(p), \text{ ceea ce este adevărat.} \end{aligned}$$

(ii) Avem că $\mathcal{M}, w_1 \not\Vdash \Diamond \Box p \rightarrow p \iff \mathcal{M}, w_1 \Vdash \Diamond \Box p$ și

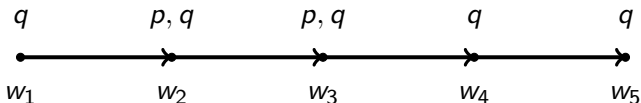
$\mathcal{M}, w_1 \not\Vdash p$. Aplicăm (i) și faptul că $w_1 \notin V(p)$.

(iii), (iv) Exercițiu.



Exemplu

Fie $\mathcal{M} = (W, R, V)$ modelul Kripke reprezentat prin:



Dem.: (continuare)

(v) Fie $w \in W$ arbitrar. Avem că $\mathcal{M}, w \Vdash \Box q \iff$ pentru orice $v \in W$, Rwv implică $\mathcal{M}, v \Vdash q \iff$ pentru orice $v \in W$, Rwv implică $v \in V(q)$, ceea ce este adevărat, deoarece $V(q) = W$.

Noțiunea de satisfacere este **internă** și **locală**. Evaluăm formulele în **interiorul** modelelor, într-o stare particulară w (**starea curentă**). În cazul operatorilor modali \Box , \Diamond nu verificăm adevărul lui φ în **toate** stările din W ci numai în acelea care sunt R -accesibile din starea curentă.

Aceasta nu este o slăbiciune a noțiunii de satisfacere, ci, dimpotrivă, ne permite o foarte mare flexibilitate. Dacă luăm $R = W \times W$, atunci toate stările sunt accesibile din w , iar dacă luăm $R = \{(v, v) \mid v \in W\}$, atunci w este singura stare accesibilă din w . Acestea sunt cazurile extreme, dar, evident, sunt multe opțiuni de explorat.

Putem să ne punem următoarele întrebări naturale: ce se întâmplă dacă impunem anumite condiții asupra lui R (de exemplu, reflexivitate, simetrie, tranzitivitate, etc.), ce impact au aceste condiții asupra necesității și posibilității, ce principii sau reguli sunt justificate de aceste condiții?



Validitatea într-un cadru Kripke este unul din conceptele cheie în logica modală.

Definiția 2.12

Fie \mathcal{F} un cadru Kripke și φ o formulă.

- ▶ φ este **validă într-o stare** w din \mathcal{F} dacă pentru orice model $\mathcal{M} = (\mathcal{F}, V)$ bazat pe \mathcal{F} , $\mathcal{M}, w \Vdash \varphi$.
- ▶ φ este **validă** în \mathcal{F} dacă este validă în orice stare w din \mathcal{F} .

Notăție: $\mathcal{F} \Vdash \varphi$

Așadar, o formulă este validă într-un cadru Kripke dacă este adevărată în orice stare din orice model bazat pe cadru.



Validitatea într-un cadru Kripke diferă în mod esențial de adevărul într-un model Kripke. Să dăm un exemplu simplu.

Dacă $\varphi \vee \psi$ este adevărată într-un model Kripke \mathcal{M} în w , atunci φ este adevărată în \mathcal{M} în w sau ψ este adevărată în \mathcal{M} în w (conform definiției satisfacției).

Pe de altă parte, dacă $\varphi \vee \psi$ este validă într-un cadru Kripke \mathcal{F} în w , nu rezultă că φ este validă în \mathcal{F} în w sau ψ este validă în \mathcal{F} în w ($p \vee \neg p$ este un contraexemplu).



Definiția 2.13

Fie \mathbf{F} o clasă de cadre și φ o formulă.

Spunem că φ este **validă în \mathbf{F}** dacă este validă în orice cadru din \mathbf{F} .

Notăție: $\mathbf{F} \models \varphi$

Definiția 2.14

Mulțimea tuturor formulelor din ML_0 care sunt valide într-o clasă de cadre \mathbf{F} se numește **logica** lui \mathbf{F} și se notează $\Lambda_{\mathbf{F}}$.



Exemplul 2.15

Pentru orice formule φ, ψ ,

$$\Diamond(\varphi \vee \psi) \rightarrow (\Diamond\varphi \vee \Diamond\psi) \quad \text{și} \quad \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$$

sunt valide în clasa tuturor cadrelor Kripke.

Dem.: Fie \mathcal{F} un cadru Kripke arbitrar, w o stare din \mathcal{F} și $\mathcal{M} = (\mathcal{F}, V)$ un model Kripke bazat pe \mathcal{F} . Trebuie să arătăm că $\mathcal{M}, w \Vdash \Diamond(\varphi \vee \psi) \rightarrow (\Diamond\varphi \vee \Diamond\psi)$.

Presupunem că $\mathcal{M}, w \Vdash \Diamond(\varphi \vee \psi)$. Atunci există $v \in W$ astfel încât Rwv și $\mathcal{M}, v \Vdash \varphi \vee \psi$. Avem două cazuri:

- ▶ $\mathcal{M}, v \Vdash \varphi$. Atunci $\mathcal{M}, w \Vdash \Diamond\varphi$, deci $\mathcal{M}, w \Vdash \Diamond\varphi \vee \Diamond\psi$.
- ▶ $\mathcal{M}, v \Vdash \psi$. Atunci $\mathcal{M}, w \Vdash \Diamond\psi$, deci $\mathcal{M}, w \Vdash \Diamond\varphi \vee \Diamond\psi$.

Lăsăm ca exercițiu demonstrația faptului că

$\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ este validă în clasa tuturor cadrelor Kripke. □



Exemplul 2.16

Formula $\Diamond\Diamond p \rightarrow \Diamond p$ nu este validă în clasa tuturor cadrelor Kripke.

Dem.: Trebuie să găsim un cadru Kripke \mathcal{F} , o stare w și un model Kripke $\mathcal{M} = (\mathcal{F}, V)$ a.î

$$\mathcal{M}, w \not\models \Diamond\Diamond p \rightarrow \Diamond p.$$

Considerăm următorul cadru: $\mathcal{F} = (W, R)$, unde

$$W = \{0, 1, 2\}, \quad R = \{(0, 1), (1, 2)\}$$

și $\mathcal{M} = (\mathcal{F}, V)$ cu $V(p) = \{2\}$. Atunci $\mathcal{M}, 0 \models \Diamond\Diamond p$, deoarece $R^2 0 2$ și $\mathcal{M}, 2 \models p$. Dar $\mathcal{M}, 0 \not\models \Diamond p$, deoarece singurul punct R -accesibil din 0 este 1 și $\mathcal{M}, 1 \not\models p$. □



Definiția 2.17

Spunem că un cadru Kripke $\mathcal{F} = (W, R)$ este **tranzitiv** dacă R este tranzitivă.

Exemplul 2.18

Pentru orice formulă φ ,

$$\Diamond\Diamond\varphi \rightarrow \Diamond\varphi$$

este validă în clasa tuturor cadrelor Kripke tranzitive.

Dem.: Fie \mathcal{F} un cadru Kripke tranzitiv, w o stare din \mathcal{F} și $\mathcal{M} = (\mathcal{F}, V)$ un model Kripke bazat pe \mathcal{F} . Presupunem că $\mathcal{M}, w \Vdash \Diamond\Diamond\varphi$. Atunci există $u, v \in W$ a.î Rwu, Ruv și $\mathcal{M}, v \Vdash \varphi$. Deoarece R este tranzitivă, rezultă că Rwv și $\mathcal{M}, v \Vdash \varphi$. Deci, $\mathcal{M}, w \Vdash \Diamond\varphi$. □



Exemplul 2.19

Verificați dacă următoarea formulă este validă în clasa tuturor cadrelor Kripke:

$$p \rightarrow \Box \Diamond p.$$

Dem.: Răspunsul este NU. Considerăm următorul cadru Kripke $\mathcal{F} = (W, R)$, unde

$$W = \{1, 2, 3\}, R = \{(1, 2), (2, 3)\}$$

și $\mathcal{M} = (\mathcal{F}, V)$ cu $V(p) = \{1\}$. Avem că $\mathcal{M}, 1 \Vdash p$, deoarece $1 \in V(p)$. Pe de altă parte,

$$\mathcal{M}, 1 \Vdash \Box \Diamond p \iff \mathcal{M}, 2 \Vdash \Diamond p \iff \mathcal{M}, 3 \Vdash p \iff 3 \in V(p),$$

ceea ce este fals.

Prin urmare, $\mathcal{M}, 1 \not\Vdash p \rightarrow \Box \Diamond p$.



Exemplul 2.20

Verificați dacă formula $\Diamond p \wedge (\Box q \vee \Box r) \rightarrow \Diamond(p \wedge (q \vee r))$ este validă în clasa tuturor cadrelor Kripke.

Dem.: Răspunsul este DA. Fie \mathcal{F} un cadru Kripke arbitrar, w o stare din \mathcal{F} și $\mathcal{M} = (\mathcal{F}, V)$ un model Kripke bazat pe \mathcal{F} . Trebuie să arătăm că $\mathcal{M}, w \Vdash \Diamond p \wedge (\Box q \vee \Box r) \rightarrow \Diamond(p \wedge (q \vee r))$.

Presupunem că $\mathcal{M}, w \Vdash \Diamond p \wedge (\Box q \vee \Box r)$, deci $\mathcal{M}, w \Vdash \Diamond p$ și ($\mathcal{M}, w \Vdash \Box q$ sau $\mathcal{M}, w \Vdash \Box r$). Fie $v \in W$ a.î.

$$(*) \quad R w v \text{ și } \mathcal{M}, v \Vdash p.$$

Dacă $\mathcal{M}, w \Vdash \Box q$, atunci pentru orice $u \in W$, $R w u$ implică $\mathcal{M}, u \Vdash q$. Luând $u = v$, obținem $\mathcal{M}, v \Vdash q$.

Dacă $\mathcal{M}, w \Vdash \Box r$, atunci pentru orice $u \in W$, $R w u$ implică $\mathcal{M}, u \Vdash r$. Luând $u = v$, obținem $\mathcal{M}, v \Vdash r$. Prin urmare,

$$(**) \quad \mathcal{M}, v \Vdash q \text{ sau } \mathcal{M}, v \Vdash r.$$

Din (*) și (**) rezultă că $R w v$ și $\mathcal{M}, v \Vdash p \wedge (q \vee r)$. Deci, $\mathcal{M}, w \Vdash \Diamond(p \wedge (q \vee r))$.



Exemplul 2.21

Verificați dacă următoarea formulă este validă în clasa tuturor cadrelor Kripke:

$$\Box\Box p \rightarrow \Box p.$$

Dem.: Răspunsul este NU. Considerăm următorul cadru Kripke $\mathcal{F} = (W, R)$, unde

$$W = \{1, 2\}, R = \{(1, 2), (2, 1)\}$$

și $\mathcal{M} = (\mathcal{F}, V)$ cu $V(p) = \{1\}$. Avem că $\mathcal{M}, 1 \Vdash \Box\Box p$
 $\iff \mathcal{M}, 2 \Vdash \Box p \iff \mathcal{M}, 1 \Vdash p \iff 1 \in V(p)$, ceea ce este adevărat.

Pe de altă parte, $\mathcal{M}, 1 \Vdash \Box p \iff \mathcal{M}, 2 \Vdash p \iff 2 \in V(p)$, ceea ce este fals.

Prin urmare, $\mathcal{M}, 1 \not\Vdash \Box\Box p \rightarrow \Box p$.





Exemplul 2.22

Verificați dacă următoarea formulă este validă în clasa tuturor cadrelor Kripke:

$$\Box p \rightarrow \Diamond p.$$

Dem.: Răspunsul este NU. Considerăm următorul cadru Kripke $\mathcal{F} = (W, R)$, unde

$$W = \{1, 2\}, R = \{(2, 2), (2, 1)\}$$

și $\mathcal{M} = (\mathcal{F}, V)$ cu V arbitrar. Deoarece nu există nicio stare $w \in W$ a.î. $R1w$, rezultă că $\mathcal{M}, 1 \Vdash \Box p$, dar $\mathcal{M}, 1 \not\Vdash \Diamond p$. □



Sintaxa

Definiția 2.23

O *logică modală normală* este o mulțime Λ de formule din ML_0 care are următoarele proprietăți:

- ▶ Λ conține următoarele *axiome*:

(*Taut*) toate tautologiile propoziționale,

(*K*) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

- ▶ Λ este închisă la următoarele reguli de deducție:

- ▶ *modus ponens (MP)*:

$$\frac{\varphi, \varphi \rightarrow \psi}{\psi}$$

Prin urmare, dacă $\varphi \in \Lambda$ și $\varphi \rightarrow \psi \in \Lambda$, atunci $\psi \in \Lambda$.

- ▶ *generalizarea (GEN)*:

$$\frac{\varphi}{\Box\varphi}$$

Prin urmare, dacă $\varphi \in \Lambda$, atunci $\Box\varphi \in \Lambda$.



Adăugăm toate tautologiile propoziționale ca axiome pentru ușurință, dar nu este necesar. Puteam să adăugăm doar un număr mic de tautologii, care le generează pe toate.

Tautologiile pot conține și modalități. De exemplu, $\Diamond\psi \vee \neg\Diamond\psi$ este tautologie, deoarece are aceeași formă cu $\varphi \vee \neg\varphi$.

Axioma (K) se mai numește și **axioma distribuției** și este importantă pentru că ne permite să transformăm $\Box(\varphi \rightarrow \psi)$ într-o implicație $\Box\varphi \rightarrow \Box\psi$, putând astfel să folosim gândirea propozițională. De exemplu, presupunem că vrem să demonstrăm $\Box\psi$ și avem deja o demonstrație care conține atât $\Box(\varphi \rightarrow \psi)$ cât și $\Box\varphi$. Atunci aplicând (K) și modus ponens, obținem $\Box\varphi \rightarrow \Box\psi$. Aplicând din nou modus ponens, rezultă $\Box\psi$.

Generalizarea “modalizează” formulele, adăugându-le \Box în față, crează noi contexte modale în care să lucrăm.



Teorema 2.24

Pentru orice clasă \mathbf{F} de cadre Kripke, logica $\Lambda_{\mathbf{F}}$ a lui \mathbf{F} este o logică modală normală.

Lema 2.25

- ▶ Colecția tuturor formulelor este o logică modală normală, numită **logica inconsistentă**.
- ▶ Dacă $\{\Lambda_i \mid i \in I\}$ este o colecție de logici modale normale, atunci $\bigcap_{i \in I} \Lambda_i$ este o logică modală normală.

Definiția 2.26

\mathbf{K} este intersecția tuturor logicilor modale normale.

Logica \mathbf{K} este cea mai mică logică modală normală, este foarte slabă. Putem obține logici mai puternice folosind ideea de a **extinde \mathbf{K} cu axiome adiționale**.

Conform Lemei 2.25, pentru orice mulțime de formule Γ , există cea mai mică logică modală normală care conține Γ .

Definiția 2.27

$K\Gamma$ este cea mai mică logică modală normală care conține Γ .
Spunem că $K\Gamma$ este *generată* de Γ sau că este *axiomatizată* de Γ .

Definiția 2.28

O *$K\Gamma$ -demonstrație* este o secvență de formule $\theta_1, \dots, \theta_n$ a.î. pentru fiecare $i \in \{1, \dots, n\}$, una din următoarele condiții este satisfăcută:

- ▶ θ_i este axiomă (adică (Taut) sau (K));
- ▶ $\theta_i \in \Gamma$;
- ▶ θ_i se obține din formule anterioare aplicând următoarele reguli de deducție: modus ponens sau generalizarea.



Definiția 2.29

Fie φ o formulă. O **$K\Gamma$ -demonstrație a lui φ** este o **$K\Gamma$ -demonstrație** $\theta_1, \dots, \theta_n$ a.î. $\theta_n = \varphi$.

Dacă φ are o **$K\Gamma$ -demonstrație**, spunem că φ este **$K\Gamma$ -demonstrabilă** și scriem $\vdash_{K\Gamma} \varphi$.

Teorema 2.30

$$K\Gamma = \{\varphi \mid \vdash_{K\Gamma} \varphi\}.$$

Luând $\Gamma = \emptyset$, obținem că $K = \{\varphi \mid \vdash_K \varphi\}$.



Exemplul 2.31

$\vdash_K \varphi \rightarrow \psi$ implică $\vdash_K \Box\varphi \rightarrow \Box\psi$.

Dem.: Prezentăm următoarea **K**-demonstrație:

- | | | |
|-----|------------------------------------------------------------------------------------------|-----------------|
| (1) | $\vdash_K \varphi \rightarrow \psi$ | ipoteză |
| (2) | $\vdash_K \Box(\varphi \rightarrow \psi)$ | (GEN): (1) |
| (3) | $\vdash_K \Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ | (K) |
| (4) | $\vdash_K \Box\varphi \rightarrow \Box\psi$ | (MP): (2), (3). |



Exemplul 2.32

$\vdash_K \varphi \rightarrow \psi$ implică $\vdash_K \Diamond\varphi \rightarrow \Diamond\psi$.

Dem.: Prezintă următoarea **K**-demonstrație:

- | | | |
|-----|----------------------------------------------------------------------------------------------------------------------|----------------------------|
| (1) | $\vdash_K \varphi \rightarrow \psi$ | ipoteză |
| (2) | $\vdash_K (\varphi \rightarrow \psi) \rightarrow (\neg\psi \rightarrow \neg\varphi)$ | (Taut) |
| (3) | $\vdash_K \neg\psi \rightarrow \neg\varphi$ | (MP): (1), (2) |
| (4) | $\vdash_K \Box\neg\psi \rightarrow \Box\neg\varphi$ | Exemplul 2.31: (3) |
| (5) | $\vdash_K (\Box\neg\psi \rightarrow \Box\neg\varphi) \rightarrow (\neg\Box\neg\varphi \rightarrow \neg\Box\neg\psi)$ | (Taut) |
| (6) | $\vdash_K \neg\Box\neg\varphi \rightarrow \neg\Box\neg\psi$ | (MP): (4), (5) |
| (7) | $\vdash_K \Diamond\varphi \rightarrow \Diamond\psi$ | definiția lui \Diamond . |



Exemplul 2.33

$$\vdash_K \Box\varphi \wedge \Box\psi \rightarrow \Box(\varphi \wedge \psi).$$

Dem.: Prezintă următoarea **K**-demonstrație:

$$(1) \quad \vdash_K \varphi \rightarrow (\psi \rightarrow (\varphi \wedge \psi)) \quad (\text{Taut})$$

$$(2) \quad \vdash_K \Box\varphi \rightarrow \Box(\psi \rightarrow (\varphi \wedge \psi)) \quad \text{Exemplul 2.31: (1)}$$

$$(3) \quad \vdash_K \Box(\psi \rightarrow (\varphi \wedge \psi)) \rightarrow (\Box\psi \rightarrow \Box(\varphi \wedge \psi)) \quad (\text{K})$$

$$(4) \quad \vdash_K \Box\varphi \rightarrow (\Box\psi \rightarrow \Box(\varphi \wedge \psi))$$

logică propozițională: (2), (3) și (MP),

$(\sigma_1 \rightarrow \sigma_2) \rightarrow ((\sigma_2 \rightarrow \sigma_3) \rightarrow (\sigma_1 \rightarrow \sigma_3))$ tautologie

$$(5) \quad \vdash_K \Box\varphi \wedge \Box\psi \rightarrow \Box(\varphi \wedge \psi)$$

logică propozițională: (4) și (MP),

$(\sigma_1 \rightarrow (\sigma_2 \rightarrow \sigma_3)) \rightarrow ((\sigma_1 \wedge \sigma_2) \rightarrow \sigma_3)$ tautologie. \square



Sintaxa și semantica



Fie \mathbf{F} o clasă de cadre Kripke și Λ o logică modală normală.

Notăție

Dacă $\varphi \in \Lambda$, spunem și că φ este **Λ -teoremă** sau **teoremă a lui Λ** și scriem $\vdash_{\Lambda} \varphi$. Dacă $\varphi \notin \Lambda$, scriem $\nvdash_{\Lambda} \varphi$.

Definiția 2.34

Λ este

- ▶ **corectă** (sound) cu privire la \mathbf{F} dacă $\Lambda \subseteq \Lambda_{\mathbf{F}}$.
- ▶ **completă** (complete) cu privire la \mathbf{F} dacă $\Lambda_{\mathbf{F}} \subseteq \Lambda$.
- ▶ Λ este corectă cu privire la \mathbf{F} ddacă pentru orice formulă φ ,
 $\vdash_{\Lambda} \varphi$ implică $\mathbf{F} \Vdash \varphi$.
- ▶ Λ este completă cu privire la \mathbf{F} ddacă pentru orice formulă φ ,
 $\mathbf{F} \Vdash \varphi$ implică $\vdash_{\Lambda} \varphi$.