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 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, s ă 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.

Logici pentru cunoaștere

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 calculelor 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 fruntile 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 noroios, copilul noroios concluzionează că trebuie să fie el.
- Niciunul dintre ceilalți copii nu știe în acest moment dacă este sau nu noroios.
- 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 $p = 3, 4, \ldots$.
- 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

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

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:
$$\psi := (\mathit{Zeus} = \mathit{Jupiter}).$$
- Fie
$$\chi := \textit{Știe(Andrei, Tată(Jupiter, Cronos))}.$$
- 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 potrivit u a în forma sa standard pentru a raționa despre cunoaștere.
- Sunt necesare formalisme alternative.

Domeniul logicilor pentru cunoaștere a fost inițiat cu publicarea, ˆn 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 caracterizata 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 $p, q, r, v, \ldots$);*
- *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.

Limbajul modal de bază

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) *Daca φ și ψ sunt formule, atunci $(\varphi \to \psi)$ 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:

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

Limbajul modal de bază

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

Conectorii propoziționali $\vee$, $\wedge$, $\leftrightarrow$ și constantele $\top$ (adevărul), $\bot$ (falsul) sunt definiți ca în logica propozițională clasică:

$$\varphi \vee \psi := \neg\varphi \to \psi \qquad \varphi \wedge \psi := \neg(\varphi \to \neg\psi)$$

$$\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi) \qquad \top := p \to p, \quad \bot := \neg\top.$$

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:

- $\Box\varphi \to \Diamond\varphi$ (ce este necesar este și posibil)
- $\varphi \to \Diamond\varphi$ (ce este, este posibil).

Ce putem spune despre formule ca $\varphi \to \Box\Diamond\varphi$ (ce este, este necesar posibil) sau $\Diamond\varphi \to \Box\Diamond\varphi$ (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

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

Presupunând că agentul nu este atotștiutor, formula $\varphi \to K\varphi$ 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 relaţie de ordine parţială pe W (adică o relaţie binară pe W care este reflexivă, antisimetrică şi tranzitivă).
- $\mathcal{F} = (W, R)$, unde R este o relaţie de echivalenţă pe W (adică o relaţie binară pe W care este reflexivă, simetrică şi tranzitivă).

Structuri relaționale

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:

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

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 $u_1, \ldots u_{n-1}$ a.î $Rwu_1, Ru_1 u_2, \ldots Ru_{n-1}v$.

Semantica

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

Rwv 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;*
- $V : PROP \to 2^W$ *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 $\mathcal{M} = (W, R, V)$.

Semantica

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 $\mathcal{M}, w \Vdash \varphi$

$$
\begin{array}{rll}
\mathcal{M}, w \Vdash p & \text{ddacă} & w \in V(p), \quad \text{unde } p \in PROP \\
\mathcal{M}, w \Vdash \neg\varphi & \text{ddacă} & \text{nu este adevărat că } \mathcal{M}, w \Vdash \varphi \\
\mathcal{M}, w \Vdash \varphi \to \psi & \text{ddacă} & \mathcal{M}, w \Vdash \varphi \text{ implică } \mathcal{M}, w \Vdash \psi \\
\mathcal{M}, w \Vdash \Box\varphi & \text{ddacă} & \text{pentru orice } v \in W, \\
 & & Rwv \text{ implică } \mathcal{M}, v \Vdash \varphi.
\end{array}
$$

Semantica

Notație

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

Pentru orice stare $w \in W$,

- $\mathcal{M}, w \Vdash \neg\varphi$ ddacă $\mathcal{M}, w \not\Vdash \varphi$.
- $\mathcal{M}, w \Vdash \varphi \to \psi$ ddacă ($\mathcal{M}, w \not\Vdash \varphi$ sau $\mathcal{M}, w \Vdash \psi$).

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,

$$\mathcal{M}, w \not\Vdash \bot$$

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

$$\mathcal{M}, w \Vdash \varphi \wedge \psi \quad \text{ddacă} \quad \mathcal{M}, w \Vdash \varphi \text{ și } \mathcal{M}, w \Vdash \psi$$

$$\mathcal{M}, w \Vdash \varphi \vee \psi \quad \text{ddacă} \quad \mathcal{M}, w \Vdash \varphi \text{ sau } \mathcal{M}, w \Vdash \psi$$

$$\mathcal{M}, w \Vdash \Diamond\varphi \quad \text{ddacă} \quad \text{există } v \in W \text{ a.î. } Rwv \text{ și } \mathcal{M}, v \Vdash \varphi.$$

Dem.: Exercițiu.

Definiția 2.11

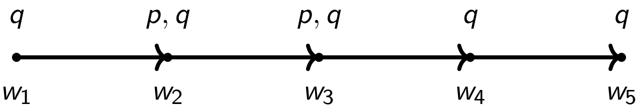
- *O formulă φ este* global adevărată *sau* universal adevărată *în $\mathcal{M}$ dacă $\mathcal{M}, w \Vdash \varphi$ pentru orice $w \in W$. Notație: $\mathcal{M} \Vdash \varphi$*
- *O formulă φ este* satisfiabilă *în $\mathcal{M}$ dacă există o stare $w \in W$ a.î. $\mathcal{M}, w \Vdash \varphi$.*

Semantica

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.

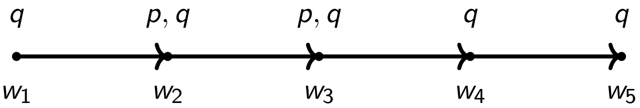
Exemplu



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) $\mathcal{M}, w_1 \nVdash \Diamond\Box p \to p$.

(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_1v și $\mathcal{M}, v \Vdash \Box p$. Luăm $v := w_2$. Cum Rw_1w_2, rămâne să demonstrăm că $\mathcal{M}, w_2 \Vdash \Box p$.

Exemplu

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

$$\begin{tikzcd} \overset{q}{w_1} \arrow[r] & \overset{p,q}{w_2} \arrow[r] & \overset{p,q}{w_3} \arrow[r] & \overset{q}{w_4} \arrow[r] & \overset{q}{w_5} \end{tikzcd}$$

Dem.: (continuare) Avem că

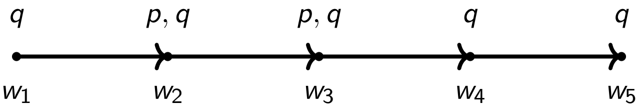
$$\begin{array}{rcl} \mathcal{M}, w_2 \Vdash \Box p & \iff & \text{pentru orice } u \in W,\ Rw_2u \text{ implică } \mathcal{M}, u \Vdash p. \\ & \iff & \mathcal{M}, w_3 \Vdash p \text{ (deoarece } w_3 \text{ este unicul } u \in W \\ & & \text{a.î. } Rw_2u) \\ & \iff & w_3 \in V(p), \text{ ceea ce este adevărat.} \end{array}$$

(ii) Avem că $\mathcal{M}, w_1 \nVdash \Diamond\Box p \to p \iff \mathcal{M}, w_1 \Vdash \Diamond\Box p$ și $\mathcal{M}, w_1 \nVdash p$. 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 $V(q) = W$.

Semantica

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 este 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 $\mathcal{F}$.*
  Notație: $\mathcal{F} \Vdash \varphi$

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

Semantica

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 $\boldsymbol{F}$ *o clasă de cadre și* φ *o formulă.*
Spunem că φ *este validă în* $\boldsymbol{F}$ *dacă este validă în orice cadru din* $\boldsymbol{F}$.
Notație: $\boldsymbol{F} \Vdash \varphi$

Definiția 2.14

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

Exemplul 2.15

Pentru orice formule φ, ψ,

$$\Diamond(\varphi \vee \psi) \to (\Diamond\varphi \vee \Diamond\psi) \quad \text{şi} \quad \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

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

$$\mathcal{M}, w \Vdash \Diamond(\varphi \vee \psi) \to (\Diamond\varphi \vee \Diamond\psi).$$

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

Exemplul 2.16

Formula $\Diamond\Diamond p \to \Diamond p$ 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

$$\mathcal{M}, w \nVdash \Diamond\Diamond p \to \Diamond p.$$

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 $\mathcal{M}, 0 \Vdash \Diamond\Diamond p$, deoarece $R^2 02$ și $\mathcal{M}, 2 \Vdash p$. Dar $\mathcal{M}, 0 \nVdash \Diamond p$, deoarece singurul punct R-accesibil din 0 este 1 și $\mathcal{M}, 1 \nVdash p$. $\square$

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,

$$\Diamond\Diamond\varphi \to \Diamond\varphi$$

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, $\mathcal{M}, w \Vdash \Diamond\varphi$. $\square$

Exemplul 2.19

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

$$p \to \Box\Diamond p.$$

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, $\mathcal{M}, 1 \nVdash p \to \Box\Diamond p$. $\square$

Exemplul 2.20

Verificaţi dacă formula $\Diamond p \wedge (\Box q \vee \Box r) \to \Diamond(p \wedge (q \vee r))$ 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 $\mathcal{M}, w \Vdash \Diamond p \wedge (\Box q \vee \Box r) \to \Diamond(p \wedge (q \vee r))$.
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.

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

Dacă $\mathcal{M}, w \Vdash \Box q$, atunci pentru orice $u \in W$, Rwu 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 $u \in W$, Rwu 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 Rwv și $\mathcal{M}, v \Vdash p \wedge (q \vee r)$. Deci, $\mathcal{M}, w \Vdash \Diamond(p \wedge (q \vee r))$. $\Box$

Exemplul 2.21

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

$$\Box\Box p \to \Box p.$$

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

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

și 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, $\mathcal{M}, 1 \nVdash \Box\Box p \to \Box p$. $\square$

Exemplul 2.22

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

$$\Box p \to \Diamond p.$$

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 $\mathcal{M}, 1 \nVdash \Diamond p$. $\Box$

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:*

$$\begin{array}{ll} (Taut) & \text{toate tautologiile propoziționale,} \\ (K) & \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi) \end{array}$$

- *Λ este închisă la următoarele reguli de deducție:*
  - *modus ponens (MP):*

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

    Prin urmare, dacă $\varphi \in \Lambda$ și $\varphi \to \psi \in \Lambda$, atunci $\psi \in \Lambda$.
  - *generalizarea (GEN):*

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

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

Sintaxa

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 $\Box(\varphi \to \psi)$ într-o implicație $\Box\varphi \to \Box\psi$, putând astfel să folosim gândirea propozițională. De exemplu, presupunem că vrem să demonstrăm $\Box\psi$ și avem deja o demonstrație care conține atât $\Box(\varphi \to \psi)$ cât și $\Box\varphi$. Atunci aplicând (K) și modus ponens, obținem $\Box\varphi \to \Box\psi$. Aplicând din nou modus ponens, rezultă $\Box\psi$.

Generalizarea "modalizează" formulele, adăugându-le $\Box$ în față, creează 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

K *este intersecția tuturor logicilor modale normale.*

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

Sintaxa

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

Definiția 2.27

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

Definiția 2.28

O $\mathbf{K}\Gamma$-demonstrație este o secvență de formule $\theta_1, \ldots, \theta_n$ a.î. pentru fiecare $i \in \{1, \ldots, n\}$, 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 ***KΓ****-demonstrație a lui φ este o* ***KΓ****-demonstrație $\theta_1, \ldots, \theta_n$ a.î. $\theta_n = \varphi$.*
Dacă φ are o ***KΓ****-demonstrație, spunem că φ este* ***KΓ****-demonstrabilă și scriem $\vdash_{\boldsymbol{K\Gamma}} \varphi$.*

Teorema 2.30

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

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

Exemplul 2.31

$$\vdash_{\boldsymbol{K}} \varphi \to \psi \text{ implică } \vdash_{\boldsymbol{K}} \Box\varphi \to \Box\psi.$$

Dem.: Prezentăm următoarea $\boldsymbol{K}$-demonstrație:

(1)	$\vdash_{\boldsymbol{K}} \varphi \to \psi$	ipoteză
(2)	$\vdash_{\boldsymbol{K}} \Box(\varphi \to \psi)$	(GEN): (1)
(3)	$\vdash_{\boldsymbol{K}} \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$	(K)
(4)	$\vdash_{\boldsymbol{K}} \Box\varphi \to \Box\psi$	(MP): (2), (3).

Exemplul 2.32

$$\vdash_{\boldsymbol{K}} \varphi \to \psi \text{ implică } \vdash_{\boldsymbol{K}} \Diamond\varphi \to \Diamond\psi.$$

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

(1)	$\vdash_{\boldsymbol{K}} \varphi \to \psi$	ipoteză
(2)	$\vdash_{\boldsymbol{K}} (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$	(Taut)
(3)	$\vdash_{\boldsymbol{K}} \neg\psi \to \neg\varphi$	(MP): (1), (2)
(4)	$\vdash_{\boldsymbol{K}} \Box\neg\psi \to \Box\neg\varphi$	Exemplul 2.31: (3)
(5)	$\vdash_{\boldsymbol{K}} (\Box\neg\psi \to \Box\neg\varphi) \to (\neg\Box\neg\varphi \to \neg\Box\neg\psi)$	(Taut)
(6)	$\vdash_{\boldsymbol{K}} \neg\Box\neg\varphi \to \neg\Box\neg\psi$	(MP): (4), (5)
(7)	$\vdash_{\boldsymbol{K}} \Diamond\varphi \to \Diamond\psi$	definiția lui $\Diamond$.

Exemplul 2.33

$$\vdash_{\boldsymbol{K}} \Box\varphi \land \Box\psi \to \Box(\varphi \land \psi).$$

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

(1) $\vdash_{\boldsymbol{K}} \varphi \to (\psi \to (\varphi \land \psi))$ (Taut)

(2) $\vdash_{\boldsymbol{K}} \Box\varphi \to \Box(\psi \to (\varphi \land \psi))$ Exemplul 2.31: (1)

(3) $\vdash_{\boldsymbol{K}} \Box(\psi \to (\varphi \land \psi)) \to (\Box\psi \to \Box(\varphi \land \psi))$ (K)

(4) $\vdash_{\boldsymbol{K}} \Box\varphi \to (\Box\psi \to \Box(\varphi \land \psi))$
logică propozițională: (2), (3) și (MP),
$(\sigma_1 \to \sigma_2) \to ((\sigma_2 \to \sigma_3) \to (\sigma_1 \to \sigma_3))$ tautologie

(5) $\vdash_{\boldsymbol{K}} \Box\varphi \land \Box\psi \to \Box(\varphi \land \psi)$
logică propozițională: (4) și (MP),
$(\sigma_1 \to (\sigma_2 \to \sigma_3)) \to ((\sigma_1 \land \sigma_2) \to \sigma_3)$ tautologie. $\Box$

Sintaxa și semantica

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

Notație

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 $\not\vdash_\Lambda \varphi$.

Definiția 2.34

Λ *este*

- *corectă (sound) cu privire la* $\boldsymbol{F}$ *dacă* $\Lambda \subseteq \Lambda_{\boldsymbol{F}}$.
- *completă (complete) cu privire la* $\boldsymbol{F}$ *dacă* $\Lambda_{\boldsymbol{F}} \subseteq \Lambda$.

- Λ este corectă cu privire la $\boldsymbol{F}$ ddacă pentru orice formulă φ,
$$\vdash_\Lambda \varphi \quad \text{implică} \quad \boldsymbol{F} \Vdash \varphi.$$
- Λ este completă cu privire la $\boldsymbol{F}$ ddacă pentru orice formulă φ,
$$\boldsymbol{F} \Vdash \varphi \quad \text{implică} \quad \vdash_\Lambda \varphi.$$

Așadar, dacă demonstrăm că o logică modală normală Λ (specificată sintactic) este atât corectă cât și completă cu privire la o clasă $\boldsymbol{F}$ de cadre Kripke, obținem o potrivire perfectă între perspectiva sintactică și cea semantică:

$$\Lambda = \Lambda_{\boldsymbol{F}}.$$

Teorema 2.35

K *este corectă și completă cu privire la clasa tuturor cadrelor Kripke.*

Fiind dată o logică modală normală $\Lambda_{\boldsymbol{F}}$ (specificată semantic), o problemă foarte importantă este găsirea unei mulțimi cât mai simple de formule Γ astfel încât $\Lambda_{\boldsymbol{F}}$ este logica generată de Γ. Spunem și că Γ axiomatizează $\Lambda_{\boldsymbol{F}}$.

LOGICI MULTIMODALE

Logici multimodale

Întreaga teorie prezentată până acum se extinde ușor la limbaje cu mai mulți operatori modali.

Fie I o mulțime nevidă.

- Limbajul multimodal ML_I constă în: o mulțime $PROP$ de propoziții atomice, $\neg$, $\rightarrow$, parantezele (,) și o mulțime de operatori modali $\{\Box_i \mid i \in I\}$.
- Formulele lui ML_I sunt definite, folosind notația Backus-Naur, astfel:

$$\varphi ::= p \mid \bot \mid (\neg\varphi) \mid (\varphi \rightarrow \psi) \mid (\Box_i \varphi),$$

  unde $p \in PROP$ și $i \in I$.
- Dualul lui $\Box_i$ se notează $\Diamond_i$ și este definit ca:

$$\Diamond_i \varphi := \neg\Box_i\neg\varphi$$

Logici multimodale

- Un cadru Kripke pentru ML_I este o structură relațională $\mathcal{F} = (W, \{R_i \mid i \in I\})$, unde R_i este relație binară pe W pentru orice $i \in I$.
- Un model Kripke pentru ML_I este o pereche $\mathcal{M} = (\mathcal{F}, V)$, unde $\mathcal{F}$ este un cadru Kripke și $V : PROP \to 2^W$ este o evaluare.
- Ultima clauză din definiția relației de satisfacere $\mathcal{M}, w \Vdash \varphi$ devine: pentru orice $i \in I$,

$$\mathcal{M}, w \Vdash \Box_i\varphi \text{ ddacă pentru orice } v \in W, R_i wv \text{ implică } \mathcal{M}, v \Vdash \varphi.$$

- Rezultă că

$$\mathcal{M}, w \Vdash \Diamond_i\varphi \text{ ddacă există } v \in W \text{ a.î. } R_i wv \text{ și } \mathcal{M}, v \Vdash \varphi.$$

- Definițiile adevărului într-un model Kripke ($\mathcal{M} \Vdash \varphi$) sau validității într-un cadru Kripke ($\mathcal{F} \Vdash \varphi$) sunt neschimbate.

Definiția 2.36
*O **logică multimodală normală** este o mulțime Λ de formule din ML_I care are următoarele proprietăți:*

- *Λ conține toate tautologiile propoziționale.*
- *Λ conține toate formulele*

$$(K) \quad \Box_i(\varphi \to \psi) \to (\Box_i\varphi \to \Box_i\psi),$$

unde φ, ψ sunt formule și $i \in I$.

- *Λ este închisă la modus ponens.*
- *Λ este închisă la generalizare: pentru orice formulă φ și orice $i \in I$,*

$$\frac{\varphi}{\Box_i\varphi}.$$

Logici multimodale

- Folosim aceeași notație, $\boldsymbol{K}$, pentru cea mai mică logică multimodală normală.
- Logica multimodală normală generată de o mulțime Γ de formule se notează tot $\boldsymbol{K\Gamma}$.
- Definim similar $\boldsymbol{K\Gamma}$-demonstrațiile și avem, de asemenea, $\boldsymbol{K\Gamma} = \{\varphi \mid \vdash_{\boldsymbol{K\Gamma}} \varphi\}$. În particular, $\boldsymbol{K} = \{\varphi \mid \vdash_{\boldsymbol{K}} \varphi\}$.
- Definițiile corectitudinii și completitudinii sunt neschimbate.

LOGICI EPISTEMICE

Logici epistemice

În logicile epistemice, limbajul multimodal este folosit pentru a raționa despre cunoaștere. Fie $n \geq 1$ și $AG = \{1, \ldots, n\}$ mulțimea agenților.

- Considerăm limbajul multimodal ML_{Ag}.
- Scriem pentru orice $i = 1, \ldots, n$, $K_i\varphi$ în loc de $\Box_i\varphi$.
- $K_i\varphi$ se citește agentul i știe că φ.
- Notăm cu $\hat{K}_i$ operatorul dual: $\hat{K}_i\varphi = \neg K_i \neg\varphi$.
- Atunci $\hat{K}_i\varphi$ se citește agentul i consideră posibil (că) φ.

- Un cadru Kripke este o structură $\mathcal{F} = (W, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, unde W este o mulțime nevidă și $\mathcal{K}_i$ este relație binară pe W pentru orice $i = 1, \ldots, n$. Așadar, scriem $\mathcal{K}_i$ în loc de R_i.
- Un model Kripke este o structură $\mathcal{M} = (\mathcal{F}, V)$, unde $V : PROP \to 2^W$.

Definiția 2.37

O *logică epistemică* *este o mulțime Λ de formule ale lui ML_{Ag} care are următoarele proprietăți:*

- *Λ conține toate tautologiile propoziționale.*
- *Λ conține toate formulele*

$$(K) \quad K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi),$$

  unde φ, ψ sunt formule și $i \in Ag$.
- *Λ este închisă la modus ponens.*
- *Λ este închisă la generalizare: pentru orice formulă φ și orice $i \in Ag$,*

$$\frac{\varphi}{K_i\varphi}.$$

Notăm cea mai mică logică epistemică tot cu ***K***.

Teorema 2.38

K *este corectă și completă privire la clasa tuturor cadrelor Kripke.*

Considerăm

$$(T) \quad K_i\varphi \to \varphi,$$

unde φ este formulă și $i \in Ag$.

Axioma (T) se numește axioma cunoașterii: Dacă un agent știe φ, atunci φ trebuie să fie adevărată. Ce este cunoscut este adevărat. Aceasta este adesea considerată ca fiind proprietatea care distinge cunoașterea de alte atitudini informaționale, cum ar fi credința.

Folosim notația ***T*** pentru logică epistemică generată de (T).

Definiția 2.39
Un cadru Kripke $\mathcal{F} = (W, \{\mathcal{K}_i \mid i \in Ag\})$ este reflexiv dacă $\mathcal{K}_i$ este reflexivă pentru orice $i \in Ag$.

Teorema 2.40
T *este corectă și completă cu privire la clasa tuturor cadrelor Kripke reflexive.*

Considerăm

$$(D) \quad \neg K_i(\varphi \wedge \neg\varphi),$$

unde φ este formulă și $i \in Ag$.

Axioma (D) este axioma consistenței: un agent nu știe atât φ cât și $\neg\varphi$. Un agent nu poate ști o contradicție.

Folosim notația ***KD*** pentru logică epistemică generată de (D).

Definiția 2.41

O relație binară R pe W se numește serială *dacă pentru orice $w \in W$ există $v \in W$ astfel încât Rwv.*

Definiția 2.42

Un cadru Kripke $\mathcal{F} = (W, \{\mathcal{K}_i \mid i \in Ag\})$ este serial *dacă $\mathcal{K}_i$ este serială pentru orice $i \in Ag$.*

Teorema 2.43

***KD** este corectă și completă cu privire la clasa tuturor cadrelor Kripke seriale.*

Considerăm

$$(4) \quad K_i\varphi \to K_iK_i\varphi,$$

unde φ este formulă și $i \in Ag$.

Axioma (4) este numită introspecție pozitivă: dacă un agent știe φ, atunci știe că știe φ. Un agent știe ce știe.

Folosim notația ***K*4** pentru logică epistemică generată de (4).

Definiția 2.44

Un cadru Kripke $\mathcal{F} = (W, \{\mathcal{K}_i \mid i \in Ag\})$ este tranzitiv dacă $\mathcal{K}_i$ este tranzitivă pentru orice $i \in Ag$.

Teorema 2.45

***K*4** *este corectă și completă cu privire la clasa tuturor cadrelor Kripke tranzitive.*

Considerăm

$$(B) \quad \varphi \to K_i \neg K_i \neg \varphi,$$

unde φ este formulă și $i \in Ag$.

Axioma (B) spune că dacă φ are loc, atunci un agent știe că nu știe $\neg\varphi$.

Folosim notația ***B*** pentru logică epistemică generată de (B).

Definiția 2.46

Un cadru Kripke $\mathcal{F} = (W, \{\mathcal{K}_i \mid i \in Ag\})$ este simetric dacă $\mathcal{K}_i$ este simetrică pentru orice $i \in Ag$.

Teorema 2.47

B *este corectă și completă cu privire la clasa tuturor cadrelor Kripke simetrice.*

Considerăm

$$(5) \quad \neg K_i\varphi \to K_i\neg K_i\varphi$$

unde φ este formulă și $i \in Ag$.

Axioma (5) este numită introspecția negativă: dacă un agent nu știe φ, atunci știe că nu știe φ. Un agent este conștient de ceea ce nu știe.

Vom folosi notația ***K*5** pentru logică epistemică generată de (5).

Definiția 2.48

O relație binară R pe W se numește euclideană *dacă pentru orice $u, v, w \in W$*

$$wRu \text{ și } wRv \text{ implică } uRv.$$

Definiția 2.49

Un cadru Kripke $\mathcal{F} = (W, \{\mathcal{K}_i \mid i \in Ag\})$ este euclidean *dacă $\mathcal{K}_i$ este Euclideană pentru orice $i \in Ag$.*

Teorema 2.50

K5 *este corectă și completă cu privire la clasa tuturor cadrelor Kripke euclideene.*

Fie $\boldsymbol{S4} = \boldsymbol{KT4}$.

Teorema 2.51

$\boldsymbol{S4}$ *este corectă și completă cu privire la clasa cadrelor Kripke ale căror relații sunt reflexive și tranzitive.*

Fie $\boldsymbol{S5} = \boldsymbol{KT4B}$. $\boldsymbol{S5}$ este considerată logica cunoașterii idealizate.

Introspecția pozitivă și negativă implică faptul că un agent are cunoaștere perfectă despre ceea ce știe și ce nu știe.

Propoziția 2.52

$\boldsymbol{S5} = \boldsymbol{KDB4} = \boldsymbol{KDB5} = \boldsymbol{KT5}$.

Teorema 2.53

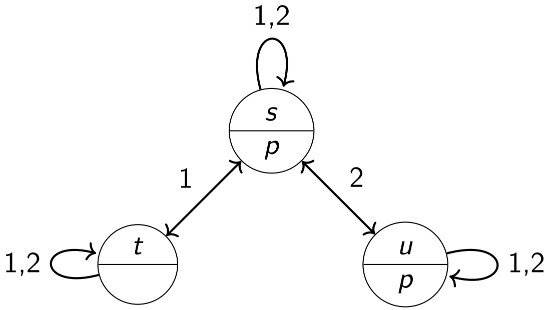
$\boldsymbol{S5}$ *este corectă și completă cu privire la clasa cadrelor Kripke ale căror relații sunt relații de echivalență.*

Logică epistemică

Un model Kripke $\mathcal{M} = (W, \mathcal{K}_1, \ldots, \mathcal{K}_n, V)$ poate fi reprezentat ca un graf etichetat:

- nodurile grafului sunt stările modelului.
- etichetăm fiecare nod cu propozițiile atomice care sunt adevărate în acel nod.
- $\mathcal{K}_i wv$ ddacă există un arc de la nodul w la nodul v, pe care îl etichetăm cu i.

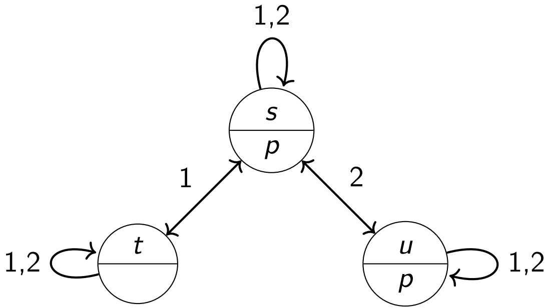
Exemplu



Avem că $Ag = \{1, 2\}$, $Prop = \{p\}$ și $\mathcal{M} = (W, R, V)$, unde

- $W = \{s, t, u\}$.
- $\mathcal{K}_1 = \{(s, s), (t, t), (u, u), (s, t), (t, s)\}$.
- $\mathcal{K}_2 = \{(s, s), (t, t), (u, u), (s, u), (u, s)\}$.
- $V(p) = \{s, u\}$.

Exemplu



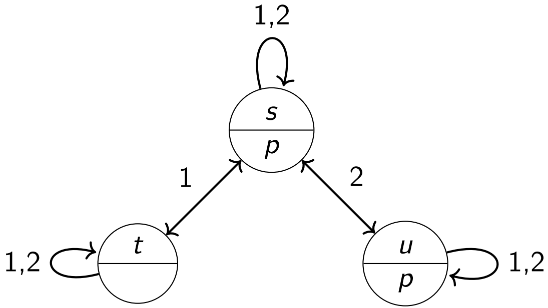
- $\mathcal{M}, s \Vdash p \wedge \neg K_1 p$.

Dem.: Avem că $s \in V(p)$, deci $\mathcal{M}, s \Vdash p$. Deoarece $\mathcal{K}_1 st$ și $\mathcal{M}, t \nVdash p$, rezultă că $\mathcal{M}, s \nVdash K_1 p$, deci $\mathcal{M}, s \Vdash \neg K_1 p$. Prin urmare, $\mathcal{M}, s \Vdash p \wedge \neg K_1 p$. □

În starea s, p este adevărată, dar agentul 1 nu știe asta, deoarece în starea s el consideră atât s cât și t posibile.

Informațiile pe care le are agentul 1 nu îi permit să distingă dacă lumea actuală este s sau t.

Exemplu

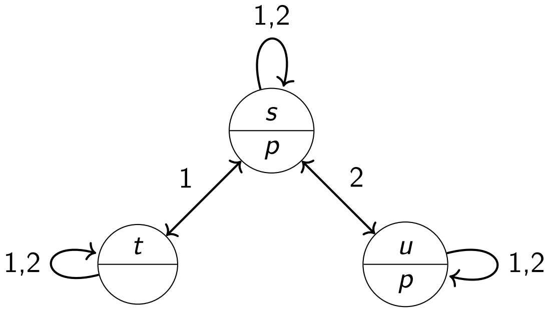


- $\mathcal{M}, s \Vdash K_2 p$.

Dem.: Avem că $\mathcal{M}, s \Vdash K_2 p$ ddacă pentru orice $v \in W$, $\mathcal{K}_2 sv$ implică $\mathcal{M}, v \Vdash p$ ddacă $\mathcal{M}, s \Vdash p$ și $\mathcal{M}, u \Vdash p$ (deoarece $\mathcal{K}_2 ss$, $\mathcal{K}_2 su$, dar nu avem că $\mathcal{K}_2 st$), ceea ce este adevărat. □

În starea s, agentul 2 știe că p este adevărată, deoarece p este satisfăcută în ambele lumi pe care le consideră posibile din s, și anume, s și t.

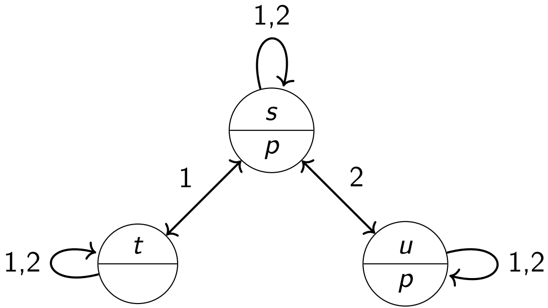
Exemplu



- $\mathcal{M}, s \Vdash \neg K_2 \neg K_1 p$.

Dem.: Avem că $\mathcal{M}, s \Vdash \neg K_2 \neg K_1 p$ ddacă $\mathcal{M}, s \nVdash K_2 \neg K_1 p$ ddacă există $v \in W$ a.î. $\mathcal{K}_2 sv$ și $\mathcal{M}, v \nVdash \neg K_1 p$ ddacă există $v \in W$ a.î. $\mathcal{K}_2 sv$ și $\mathcal{M}, v \Vdash K_1 p$. Luăm $v := u$. Atunci $\mathcal{K}_2 su$ și $\mathcal{M}, u \Vdash K_1 p$, deoarece $\mathcal{M}, u \Vdash p$ și $\mathcal{K}_1 uw$ ddacă $w = u$. $\square$

Exemplu



- $\mathcal{M}, s \Vdash \neg K_2 \neg K_1 p$.

Chiar dacă agentul 2 știe că p este adevărată în starea s, el nu știe că agentul 1 nu știe acest lucru. De ce? Pentru că într-o lume pe care agentul 2 o consideră posibilă, și anume u, agentul 1 știe că p are loc, dar în altă lume posibilă, și anume s, agentul 1 nu știe acest lucru.

Fie $\mathcal{F} = (W, \mathcal{K}_1, \ldots, \mathcal{K}_n)$ un cadru Kripke, w o stare din $\mathcal{F}$ și $\mathcal{M} = (\mathcal{F}, V)$ un model Kripke bazat pe $\mathcal{F}$, φ o formulă și $i \in Ag$.

Dacă $\mathcal{F}$ este reflexiv, atunci

$$\mathcal{M}, w \Vdash K_i\varphi \to \varphi.$$

Dem.: Presupunem că $\mathcal{M}, w \Vdash K_i\varphi$. Deoarece $\mathcal{K}_i ww$, rezultă că $\mathcal{M}, w \Vdash \varphi$. □

Dacă $\mathcal{F}$ este serial, atunci

$$\mathcal{M}, w \Vdash \neg K_i(\varphi \wedge \neg\varphi).$$

Dem.: Trebuie să arătăm că $\mathcal{M}, w \nVdash K_i(\varphi \wedge \neg\varphi)$. Deoarece $\mathcal{K}_i$ este serială, există $v \in W$ a.î. $\mathcal{K}_i wv$. Pentru orice astfel de v avem că $\mathcal{M}, v \nVdash \varphi \wedge \neg\varphi$, deoarece nu putem avea că $\mathcal{M}, v \Vdash \varphi$ și $\mathcal{M}, v \nVdash \varphi$. □

Dacă $\mathcal{F}$ este tranzitiv, atunci

$$\mathcal{M}, w \Vdash K_i\varphi \to K_iK_i\varphi.$$

Dem.: Exercițiu.

Dacă $\mathcal{F}$ este simetric, atunci

$$\mathcal{M}, w \Vdash \varphi \to K_i\neg K_i\neg\varphi.$$

Dem.: Exercițiu.

Dacă $\mathcal{F}$ este euclidean, atunci

$$\mathcal{M}, w \Vdash \neg K_i\varphi \to K_i\neg K_i\varphi.$$

Dem.: Presupunem că $\mathcal{M}, w \Vdash \neg K_i\varphi$, deci că $\mathcal{M}, w \nVdash K_i\varphi$. Rezultă că există $v \in W$ a.î. $\mathcal{K}_i wv$ și $\mathcal{M}, v \nVdash \varphi$. Fie $u \in W$ a.î. $\mathcal{K}_i wu$. Trebuie să arătăm că $\mathcal{M}, u \Vdash \neg K_i\varphi$, deci că $\mathcal{M}, u \nVdash K_i\varphi$. Din $\mathcal{K}_i wu$, $\mathcal{K}_i wv$ și $\mathcal{K}_i$ este euclideană, rezultă că $\mathcal{K}_i uv$. Deoarece $\mathcal{K}_i uv$ și $\mathcal{M}, v \nVdash \varphi$, obținem că $\mathcal{M}, u \nVdash K_i\varphi$. $\square$

MODELUL CU PARTIȚII PENTRU CUNOAȘTERE

Modelul cu partiții pentru cunoaștere

Modelele cu partiții pentru cunoaștere (*partition models of knowledge*) sunt definite în

Yoav Shoham, Kevin Leyton-Brown, Multiagents Systems, Cambridge University Press, 2009.

Reamintim că $n \geq 1$ și $AG = \{1, \ldots, n\}$ este mulțimea agenților.

Definiția 2.54

Un cadru cu partiții (partition frame) este un tuplu $\mathcal{P}_F = (W, I_1, \ldots, I_n)$*, unde*

- W *este o mulțime nevidă de lumi posibile.*
- *Pentru orice* $i = 1, \ldots, n$*,* I_i *este o partiție a lui* W*.*

Ideea este că I_i partiționează W în mulțimi de lumi posibile care sunt imposibil de distins (*indistinguishable*) din punctul de vedere al agentului i.

Fie A o mulțime nevidă.

Recapitulare: O partiție a lui A este o familie $(A_j)_{j\in J}$ de submulțimi nevide ale lui A care verifică proprietățile:

$$A = \bigcup_{j\in J} A_j \text{ și } A_j \cap A_k = \emptyset \text{ pentru orice } j \neq k.$$

Recapitulare: Există o bijecție între mulțimea relațiilor de echivalență pe A si mulțimea partițiilor lui A:

- $(A_j)_{j\in J}$ partiție a lui $A \mapsto$ relația de echivalență pe A definită prin: $x \sim y \Leftrightarrow$ există $j \in J$ astfel încât $x, y \in A_j$.
- $\sim$ relație de echivalență pe $A \mapsto$ partiția care constă în toate clasele de echivalență diferite ale lui $\sim$.

Modelul cu partiții pentru cunoaștere

Fie $i = 1, \ldots, n$,

- Notăm $\mathcal{K}_i$ relația de echivalență corespunzătoare partiției I_i.
- Pentru orice lume $w \in W$, fie $I_i(w)$ clasa de echivalență a lui w în relația $\mathcal{K}_i$. $I_i(w)$ este mulțimea lumilor posibile pe care agentul i nu le poate distinge de w.
- Pentru orice lumi $v, w \in W$, $v \in I_i(w)$ ddacă $\mathcal{K}_i wv$ ddacă agentul i nu poate distinge v de w.

$\mathcal{F} = (W, \mathcal{K}_1, \ldots, \mathcal{K}_n)$ este un cadru Kripke pentru logica epistemică $\boldsymbol{S5}$.

Cadru cu partiții = cadru Kripke pentru logica epistemică $\boldsymbol{S5}$

Definiția 2.55

Un model cu partiții *(partition model) peste un limbaj* Σ *este un tuplu* $\mathcal{P}_M = (\mathcal{P}_F, \pi)$*, unde*

- $\mathcal{P}_F = (W, I_1, \ldots, I_n)$ *este un cadru cu partiții.*
- $\pi : \Sigma \to 2^W$ *este o funcție de interpretare.*

Pentru orice afirmație $\varphi \in \Sigma$, gândim $\pi(\varphi)$ ca fiind mulțimea lumilor posibile din modelul cu partiții $\mathcal{P}_M$ unde φ este satisfăcută.

Putem lua, de exemplu, Σ ca fiind o mulțime de propoziții atomice sau o mulțime de formule din logica propozițională.

Modelul cu partiții pentru cunoaștere

Folosim notația $K_i\varphi$ pentru "agentul i știe că φ".

Definim în continuare ce înseamnă că o afirmație este adevărată într-un model cu partiții.

Definiția 2.56

Fie $\mathcal{P}_M = (W, I_1, \ldots, I_n, \pi)$ un model cu partiții peste Σ și $w \in W$. Definim relația $\models$ (logical entailment) astfel:

- *Pentru orice $\varphi \in \Sigma$, spunem că $\mathcal{P}_M, w \models \varphi$ ddacă $w \in \pi(\varphi)$.*
- *$\mathcal{P}_M, w \models K_i\varphi$ ddacă pentru orice lume $v \in W$, $v \in I_i(w)$ implică $\mathcal{P}_M, v \models \varphi$.*

Model cu partiții = Kripke pentru logica epistemică ***S5***.

Putem raționa riguros despre cunoaștere în termeni de modele cu partiții, deci folosind logica epistemică.

Modelul cu partiții pentru cunoaștere

Aplicăm modelul cu partiții pentru cunoaștere la puzzle-ul cu copiii noroioși (*Muddy children puzzle*).

- Considerăm cazul $n = k = 2$ (doi copii, amândoi noroioși).
- Avem două propoziții atomice: muddy1 și muddy2.
- Există patru lumi posibile:

$$\begin{aligned} w_1 &: \quad muddy1 \wedge muddy2 \quad \text{(lumea reală)} \\ w_2 &: \quad muddy1 \wedge \neg muddy2 \\ w_3 &: \quad \neg muddy1 \wedge muddy2 \\ w_4 &: \quad \neg muddy1 \wedge \neg muddy2. \end{aligned}$$

- Prin urmare, $\pi(muddy1) = \{w_1, w_2\}$ și $\pi(muddy2) = \{w_1, w_3\}$.
- Sunt două partiții I_1 și I_2.

Modelul cu partiții pentru cunoaștere

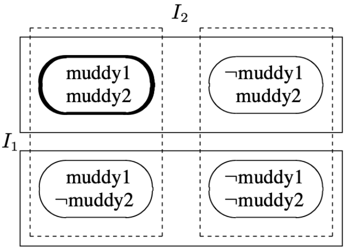
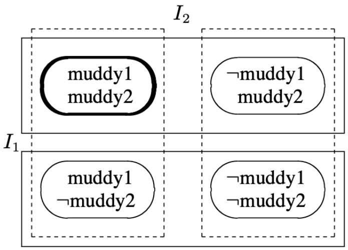


Figura 1: Modelul cu partiții după ce copiii se văd

- Ovalele ilustrează cele patru lumi posibile. Ovalul îngroșat indică adevărata stare a lumii.
- Dreptunghiurile solide reprezintă clasele de echivalență din I_1:
  $I_1(w_1) = I_1(w_3) = \{w_1, w_3\}, \quad I_1(w_2) = I_1(w_4) = \{w_2, w_4\}$
- Dreptunghiurile punctate reprezintă clasele de echivalență din I_2:
  $I_2(w_1) = I_2(w_2) = \{w_1, w_2\}, \quad I_2(w_3) = I_2(w_4) = \{w_3, w_4\}$.

Modelul cu partiții pentru cunoaștere



În lumea reală w_1,

- $K_1 muddy2$ și $K_2 muddy1$ sunt adevărate.
- $K_1 muddy1$ nu este adevărată.
- $K_2 muddy2$ nu este adevărată.

Modelul cu partiții pentru cunoaștere

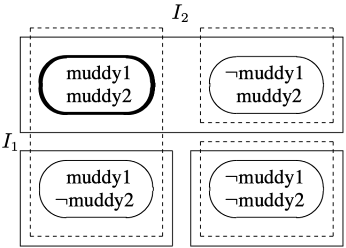


Figura 2: Modelul cu partiții după anunțul tatălui

- Lumea în care niciun copil nu este murdar este eliminată.
  Starea cunoașterii este prezentată în Figura 2:
  $I_1(w_1) = I_1(w_3) = \{w_1, w_3\}, \quad I_1(w_2) = \{w_2\}, \quad I_1(w_4) = \{w_4\}$
  $I_2(w_1) = I_2(w_2) = \{w_1, w_2\}, \quad I_2(w_3) = \{w_3\}, \quad I_2(w_4) = \{w_4\}.$
- În lumea reală w_1,
  - $K_1 muddy1$ nu este adevărată.
  - $K_2 muddy2$ nu este adevărată.

Modelul cu partiții pentru cunoaștere

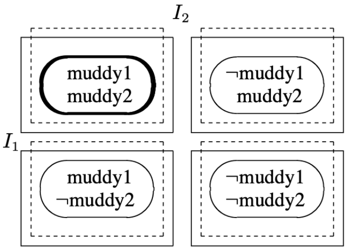


Figura 3: Modelul cu partiții final

- Fiecare copil observă că celălalt copil nu ridică mâna după anunțul tatălui. Prin urmare, fiecare își dă seama că trebuie să fie noroios.
- Starea cunoașterii este prezentată în Figura 3:
  $I_k(w_i) = \{w_i\}$ for all $k = 1, 2,\ i = 1, \ldots, 4$.
- Atât $K_1 muddy1$ cât și $K_2 muddy2$ sunt satisfăcute în lumea reală w_1.