

Seriál – neklasické logiky

I. Klasická výroková logika

Nediv se, že seriál o neklasických logikách začínáme povídáním o logice klasické. Za prvé nám to umožní lépe pochopit, co je na neklasických logikách nového, zajímavého a neobvyklého; za druhé Ti dobrá znalost klasické logiky může pomoci všude tam, kde je třeba se jasně a jednoznačně vyjadřovat, například při řešení matematických úloh, ať už v našem semináři nebo ve škole.¹

Části označené hvězdičkou se Ti možná při prvním čtení budou zdát celkem složité, můžeš je ale bez obav přeskóčit a vrátit se k nim později.

Pohádka

Byl jednou jeden Matematik a ten si přál, aby nikdo nikdy nepochyboval o pravdivosti toho, co řekl. Jeho kamarádi Literát a Filozof, ti se v různých slovních potyčkách a půtkách cítili jako doma, ba dokonce někdy to vypadalo, že kdyby s nimi všichni souhlasili, vyšla by všechna jejich snaha o originalitu naprázdno. Ne tak Matematik. Sedával celé dny doma a přemýšlel, jak to udělat, aby si ho škodolibí školáci konečně přestali dobírat, je-li součet úhlů v trojúhelníku skutečně 180° a jestli i při teplotě -273°C platí, že $1 + 1 = 2$.

Jednoho dne navštívil Matematika Hráč a povídá: „Řekni nějaké tvrzení.“

Matematik chvíli přemýšlel, no nic rozumného ho zrovna nenapadlo, ale na dvoře si zrovna zpívalo nějaké děvčátko, tak zopakoval, co slyšel: „Kočka leze dírou, pes oknem.“

Hráč nejprve vykoul z okna, a pak povídá: „Tak ty tvrdíš, že kočka leze dírou a že pes leze oknem, rozumím tomu správně? No dobrá, kočku vidím tamhle vylézat dírou ze sklepa. Dovol mi vybrat si druhou polovinu tvé věty. Pes leze oknem. Můžeš to prokázat?“

Matematik samozřejmě nemohl, a tak musel uznat, že řekl něco, co není pravda. Zadoufal ale, že se mu podaří podobně nacytat Hráče, a tak povídá: „Teď zkus říct nějaké tvrzení ty.“

Hráč ještě jednou mrkl z okna, a pak pravil: „Jsem čínský velvyslanec nebo u vás ve dvoře stojí docela nový mercedes.“

Matematik užasl a vyhrkl: „To mi teda řekni, co z toho je pravda.“

Hráč ukázal z okna a říká: „Taky jsem nevěděl, že máte tak bohaté sousedy!“

Na dvoře totiž stál docela nový zářivě žlutý mercedes.

Hintikkova logická hra*

Pak se Hráč zamračil a říká: „Myslel jsem, že budeš v téhle hře docela dobrý.“

„V jaké hře?“ divil se matematik.

„No v té, kterou jsme zrovna hráli. Vymyslel ji slavný finský logik Jaako Hintikka už někdy v 60. letech 20. století a ty ji ještě neznáš?“

První hráč – říká se mu proponent – řekne nějaké tvrzení. Úkolem druhého hráče, tomu se říká oponent, je ukázat, že tvrzení prvního hráče není pravdivé. Musí při tom ale postupovat přesně podle pravidel.

Když to tvrzení je tvaru „**A a B**“, tak si oponent může vybrat, jestli budou dál hrát s A nebo s B. Takže když jsi ty v roli proponenta řekl, že kočka leze dírou a pes oknem, mohl jsem si vybrat, že se dál chci bavit o větě „Pes leze oknem.“

¹Možná jsi už o logice něco slyšel ve škole v hodinách matematicky nebo filozofie. V tom případě můžeš přeskóčit část nadepsanou „Výroková logika – povídání o významu“. V té se navíc oproti tomu, co se běžně učí ve školách, dočteš pouze něco málo o podivnosti implikace.

Když to tvrzení je tvaru „**A nebo B**“, tak si vybrat může proponent. Takže když jsem já v roli proponenta řekl, že jsem čínský velvyslanec nebo u vás ve dvoře stojí docela nový mercedes, mohl jsem si vybrat zase já, k čemuž jsi mě ostatně sám pobídl větou „To mi teda řekni, co z toho je pravda.“

Když to tvrzení je tvaru „**ne A**“ nebo „**není pravda, že A**“, tak si prohazují role a dál se hraje s tvrzením A. Třeba kdybys ty řekl „Dnes neprší.“, hráli bychom dál jako kdybych já řekl „Dnes prší.“

Nakonec se dojde k nějaké docela jednoduché větě, o které je i bez hraní jasné, jestli je pravdivá – v tom případě vyhrává ten, kdo je zrovna proponentem, nebo nepravdivá – to vyhraje oponent. Rozumíš?²

Cvičení 1. Zkus si rozmyslet, jak by asi mohla být formulována pravidla pro výroky tvaru *Pro každé x platí, že ...* a tvaru *Existuje nějaké takové x, že ...*

Napiš průběh hry, ve které proponent začne následujícími tvrzeními:

- (a) Pro každé přirozené číslo n platí: jestliže je n dělitelné číslem 6, je dělitelné číslem 2 a také je dělitelné číslem 3.
- (b) Existuje přirozené číslo, které je dělitelné 2 a není dělitelné 3.

Výroková logika – povídání o významu²

Matematik rozuměl a hnedka ho napadla důležitá věc: je důležité mít nějaká pravidla, která určují, co tím myslíme, když třeba řekneme, že platí A a B , přesně jako to říkají pravidla té hry. Pochopil, že prvním krokem k tomu, aby nikdo nemohl pochybovat o správnosti toho, co říká, je každému vysvětlit, co tím myslí, když používá slovíčka jako *a*, *nebo*, *jestliže*, *pak*. Což o to, se slůvky *a* a *nebo* až tak veliký problém nebude, ale jeho oblíbené *jestliže A, pak B*, to bude oříšek.

Přemýšlel a přemýšlel a nakonec sepsal tohle:

Vědu zabývající se vztahy mezi pravdivostí složené věty a pravdivostí jejích částí nazvu *logika*. Půjde v ní o to, co nejlépe zachytit a popsat *vyplývání*, tedy to, jak jedna pravda vynucuje jinou pravdu.

Přesně tak, jako to známe z češtiny, se i v logice složené výroky (v češtině jim říkáme souvětí) vytvářejí z jednodušších pomocí spojek. Přitom logické spojky vyjadřují určité vztahy mezi významem jednodušších výroků a významem celku.

Jsou-li oba výroky, které spojuje spojka **a**, pravdivé, je i složený výrok se spojkou **a** pravdivý, jinak je složený výrok nepravdivý.³

²Jak v lingvistice, tak v logice (není divu, že mají aspoň něco společného, vždyť se obě zabývají jazykem, i když z úplně jiného úhlu pohledu), se otázkami souvisejícími s významem vět a slov zabývá *sémantika*. V logice je pro nás na významu nějaké věty nejdůležitější to, zda je pravdivá nebo ne, a jak její pravdivost souvisí s pravdivostí ostatních výroků. Proto logiky, na rozdíl od lingvistů, zajímají jen *výroky*, tedy věty, které jsou buďto pravdivé nebo nepravdivé.

³O tom, že ne vždy slůvko „a“ označuje konjunktci výroků, svědčí následující příklady:

- (a) Firma Novák a syn.
- (b) Přišel domů a zjistil, že je vykradeno. (Pro konjunktci totiž musí platit „zřejmě“ formule $(A \wedge B) \Rightarrow (B \wedge A)$. Ovšem věta „Zjistil, že je vykradeno, a přišel domů.“ říká něco trochu jiného než původní věta.)

Další příklady vět, ve kterých „a“ neoznačuje konjunktci, najdeš ve cvičeních.

K tomu, aby byl výrok se spojkou **nebo** pravdivý stačí, aby byl pravdivý alespoň jeden z obou spojovaných výroků.

Vztah vyplývání mezi dvěma výroky zachycuje implikace, kterou vyjadřují výrazy jako **Jestliže A, pak B**. **Kdykoli A, potom B**. **Nechť A. Pak B**. **A implikuje B**.

Implikace je nepravdivá, pokud předpoklad A je pravdivý, ale závěr B je nepravdivý; jinak je pravdivá.⁴

Ekvivalencí mezi dvěma výroky rozumím tvrzení, že **A právě tehdy, když B**. Někdy pro ni také používáme obrat „**A tehdy a jen tehdy, když B**.“ Ekvivalence je pravdivá, jsou-li A a B buďto oba pravdivé, nebo oba nepravdivé. Lze to říct i jinak: ekvivalence je pravdivá, pokud A implikuje B a B implikuje A.

„**Není pravda, že**“ je také logická spojka, ačkoli to není spojka z hlediska jazykovědy. Tato spojka nespojuje dva výroky, ale z jednoho výroku vytváří složitější výrok. Přitom z pravdivého výroku dělá výrok nepravdivý a naopak.

Pro přehlednost můžeme význam těchto spojek zachytit tabulkou. Je-li výrok pravdivý, budeme říkat, že jeho *pravdivostní hodnota* je 1, je-li nepravdivý, je jeho pravdivostní hodnota 0.

		ne A	A a B	A nebo B	jestli A, pak B	A právě když B
		$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
A	B	negace	konjunkce	disjunkce	implikace	ekvivalence
1	1	0	1	1	1	1
1	0	0	0	1	0	0
0	1	1	0	1	1	0
0	0	1	0	0	1	1

Cvičení 2. Obraty jako „**A nebo B, ne však obojí**.“, „**Buď A, nebo B**.“, „**Právě jedna z možností A, B je pravdivá**.“ a podobně neoznačují logickou spojku *nebo*, ale jinou, která se často nazývá „**vyučovací nebo**“. Jako cvičení si pro ni zkus napsat tabulku.

Cvičení 3. Jak rozumíš českému slůvku „leda“? Zahraj si na chvíli na logika a zkus nalézt tabulku, která by zachytila pravdivostní hodnotu výroků „A, leda že by B.“ a „Jestliže A, pak B, leda že by C.“ („A, i baže B. Ak A, tak B, i baže C.“)

Domníváš se, že tyto tabulku lze navrhnout jediným možným způsobem? Lze tabulku pro „Jestliže A, pak B, leda že by C.“ vyplnit pouze na základě znalosti tabulky pro „A, ledaže by B.“?

Na základě tabulek, které navrhneš, nalezni výrok obsahující pouze obvyklé výrokové spojky, který je logicky ekvivalentní s výrokem „A, leda že by B.“

Popiš všechny situace, ve kterých jsou pravdivé následující výroky:

- Zůstanu doma s dětmi, ledaže bys mě zastoupil.
- Pes, který štěká, nekouše – leda že by štěkal. (Nebo přesněji: Jestliže pes štěká, nekouše – leda že by štěkal.)

⁴To se Ti může na první pohled zdát divné, ale zkus si představit následující situaci: kamarádka Ti řekne „Nebude-li pršet, půjdu do kina.“ V případě, že pršet nebude, je jednoduché o pravdivosti jejího tvrzení rozhodnout. Pokud kamarádka do kina půjde, mluvila pravdu. Pokud do kina nepůjde, může se na ni právem zlobit. Na případ, že pršet bude, se její výrok ale vůbec nevztahuje; jak tedy rozhodnout, mluvila-li pravdu? Asi by se Ti ale zdálo nefér kamarádku osočit z toho, že lhala. Jenže logika požaduje, aby každá věta, kterou se zabýváme, byla buďto pravdivá, nebo nepravdivá – není-li nepravdivá, musí tedy být pravdivá.

Pozorně si prohlédni tabulku implikace! Implikace je nepravdivá pouze tehdy, když výrok A je pravdivý a výrok B nikoli.

Když se rozhodneme výrokům tvaru „**Jestliže A, pak B.**“ rozumět tak, jak nám radí tabulka, budou pravdivé i následující překvapivé věty:

Jestliže 59876542665 : 123456789 je celé číslo, tak $1+1=2$.

Jestliže $1+1=2$, tak tohle je seriál v matematickém korespondenčním semináři a sluníčko je žluté.

Jestliže jsem čínská princezna, tak $1+1=2$.

Jestliže jsem čínská princezna, tak je dneska pátek třináctého.

Zkratka a dobře, libovolný pravdivý výrok je implikován libovolným jiným výrokem, který neježe s ním vůbec nemusí souviset (jako v prvních dvou příkladech), ale může dokonce být i nepravdivý! Navíc, libovolný nepravdivý výrok implikuje úplně cokoli, ať už to je pravda (první příklad s čínskou princeznou), nebo není (druhý příklad s čínskou princeznou).

Ač se to může zdát překvapující, je skutečně těžké rozumně matematicky či formálně popsat vztah mezi A a B, který normálně vyjadřujeme slovy „**Jestliže A, tak B.**“ Řešení, které našel náš matematik, a kterého se skutečně drží většina matematiků na celém světě, je krásné a jednoduché v tom, že ho lze popsat pomocí nul a jedniček tak, jak jsme to udělali v tabulce. Znamená to, že **pravdivostní hodnota složeného výroku závisí pouze na pravdivostní hodnotách jednodušších výroků, které obsahuje.**⁵ Tomu, jak rozumíme implikaci v normální lidské řeči, to moc neodpovídá: když řeknu „Jestli dostanu z té písemky pětku, budu doma bita,“ budou tomu moji spolužáci rozumět tak, že budu bita kvůli té pětce z písemky, a ne třeba proto, že jsem rozbila vzácnou vázu po dědečkovi.

Přesto má tento způsob chápání logických spojek nesporné výhody:

Pravdivost nějakého složitějšího tvrzení můžeme ověřovat tabulkou. Některá tvrzení, říká se jim *tautologie*, budou ohodnocena jedničkou bez ohledu na to, jaké pravdivostní hodnoty mají jejich části.

Podívejme se třeba na to, jak vypadá tabulka pro tvrzení tvaru $(A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A))$. Postupně budeme zjišťovat hodnoty složitějších a složitějších výroků v jednotlivých řádcích tabulky:

A	B	$A \Rightarrow B$	$B \Rightarrow A$	$(A \Rightarrow B) \wedge (B \Rightarrow A)$	$A \Leftrightarrow B$	$(A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A))$
1	1	1	1	1	1	1
1	0	0	1	0	0	1
0	1	1	0	0	0	1
0	0	1	1	1	1	1

Cvičení 4. Zkus si vyplnit tabulku třeba pro tvrzení tvaru $A \vee \neg A$ a pro tvrzení tvaru $(A \wedge (A \Rightarrow B)) \Rightarrow B$.

Jiná tvrzení, těm se říká *kontradikce*, budou vždy mít pravdivostní hodnotu 0.

⁵*Téhle vlastnosti se říká, že spojky výrokové logiky jsou *extenzionální*. Až budeme mluvit o modálních logikách, všimni si, že spojka „**možná (=může být pravda, že)**“ není extenzionální (takovým spojkám se pak říká intenzionální). Například věta „*Možná zítra bude pršet.*“ je pravdivá, bez ohledu na to, zda věta „*Zítra bude pršet.*“ je pravdivá. Pravdivost složeného výroku závisí nejen na pravdivosti jeho částí, ale i na tom, jakým způsobem jsou tyto pravdivosti popsány – nejde jen o to, zda hodnota je 1 nebo 0, ale i o to, jak se k tomu dojde.

Nepotřebujeme všechny spojky – ve skutečnosti stačí negace a implikace, nebo negace a konjunkce. Ostatním spojkám lze rozumět jako zkratkám za složitější výroky napsané pomocí vybraných spojek.

Cvičení 5. Dejme tomu, že se Ti líbí spojka \neg a spojka \wedge , zato \vee se Ti nelíbí (plete se s písmenem V). Ověř tabulkou, že $A \vee B$ má ve všech řádcích stejné hodnoty jako $\neg(\neg A \wedge \neg B)$! To znamená: Dokaž, že $(A \vee B) \Leftrightarrow \neg(\neg A \wedge \neg B)$ je tautologie.

Cvičení 6. Ověř, že A právě tehdy, když B , vlastně znamená A a B mají vždycky stejné hodnoty. Tedy jinak řečeno, ukaž, že $A \Leftrightarrow B$ má stejné pravdivostní hodnoty jako $(A \wedge B) \vee (\neg A \wedge \neg B)$!

Cvičení 7. Najdi formuli, kterou nelze ekvivalentně zapsat pouze pomocí spojek $\wedge, \vee, \Rightarrow, a \Leftrightarrow$.

Cvičení 8. Podívej se na tabulku výrokové spojky, jejíž význam je „ani A, ani B“.

A	B	$A \otimes B$
1	1	0
1	0	0
0	1	0
0	0	1

Ukaž, že pomocí této jediné spojky lze vyjádřit všechny spojky výrokové logiky, tedy najdi formule, které obsahují pouze spojku \otimes a jsou ekvivalentní s $\neg A, A \wedge B, A \Rightarrow B, A \vee B, A \Leftrightarrow B$.

Toho, že nepotřebujeme všechny spojky, se hodně využívá v informatice. Každému počítačovému programu totiž lze rozumět tak, že jen počítá pravdivostní hodnoty nějakých složitých formulí. K tomu, aby to mohl udělat, potřebuje v sobě mít zabudované součástky, které umí spočítat pravdivostní hodnotu negace, konjunkce, disjunkce, implikace a ekvivalence z pravdivostních hodnot jejich částí. Může se ale stát, že je o hodně levnější nebo jednodušší vyrobit součástku, která počítá konjunkci, než třeba součástku, která počítá implikaci ...

Výroková logika – povídání o formě⁶

Vraťme se k našemu Matematikovi. Když přišel na to, že pravdivostní hodnoty složených výroků se dají určovat pomocí tabulek, a že existují tautologie, tedy tvrzení pravdivá bez ohledu na pravdivost jednotlivých částí, začalo ho zajímat, jestli existuje nějaký způsob, jak napsat počítačový program, který by je postupně vypsal všechny. Brzy mu došlo, že těch vždy pravdivých tvrzení je nekonečně mnoho, takže to nejlepší, co se mu může podařit, je najít program, který bude postupně vypisovat jedno za druhým a na každé někdy přijde řada (i když on sám se toho možná už nedožije, a když to bude nějaká hodně složitá tautologie, nedožije se toho ani jeho pravnuček).

Nakonec takový program objevil. My si ho zde nebudeme ukazovat, ale ukážeme si něco, s pomocí čeho už by se dal napsat. V logice se takovýmhle schémátům pro programy, které vypisují jednu tautologii za druhou, říká *kalkuly*. Kalkuly a počítačové programy vůbec samozřejmě

⁶Tuhle část bychom také mohli nadepsat cizím slovem *syntaxe*. Syntaxe se zabývá tím, jak vypadá správně utvořený výrok, jaké musí mít části, tak jako věta musí mít podmět a přísudek a souvětí musí být správně rozděleno čárkami. Zajímavé na tom je, že při zkoumání správně utvořených složených výroků můžeme zapomenout na to, jaký mají význam. Dále se syntaxe zabývá vším, co lze zkoumat pouze na základě tvaru (formy) jednotlivých tvrzení, aniž bychom znali jejich konkrétní význam.

nevědí nic o tom, jestli venku prší, a dokonce ani o matematických pojmech, číslech, trojúhelnících, zkratka ničemu nerozumí. Proto jsou čistě formální – takový kalkul pracuje tak, že vezme nějakou řádku znaků, a udělá s ní nějakou jednoduchou úpravu podle předem daného pravidla. To jen člověk, který se na to dívá z venku, ví, že ta úprava znamenala třeba to, že z výroku $A \wedge (A \Rightarrow B)$ vyplývá výrok B . Kalkul sám nic takového „neví“.

Kalkul také neobsahuje ani nevytváří úplné věty, ale jen zjednodušená schémata, podle kterých lze z jednoduchých výroků vytvořit složený výrok. Takovým schématům se říká *formule*. Formule je taková řádka znaků, která by byla výrokiem, kdybychom za písmena v ní obsažená dosadili nějaké jednoduché výroky. Třeba $A \wedge (A \Rightarrow B)$ je formule, a dosazením „Prší.“ za A a „Je mokro.“ za B dostaneme výrok „Prší a když prší, tak je mokro.“

Každý kalkul obsahuje nějaké axiomy, které říkají, že na začátku smíme vzít libovolnou řádku znaků takového a takového tvaru. V kalkulu výrokové logiky to jsou tři axiomy

- (A1) $A \Rightarrow (B \Rightarrow A)$
 (A2) $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$
 (A3) $(\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B)$

Nenech se překvapit tím, že se zde vyskytují jen dvě spojky. Mohli bychom, samozřejmě, napsat i axiomy obsahující ostatní spojky, my se ale spokojíme s tím, že víme, že ostatní spojky jsou jen zkratkami za složité formule obsahující jen \Rightarrow a \neg .⁷

To, že náš kalkul obsahuje axiomy (A1), (A2) a (A3) znamená, že příslušný program může do seznamu tautologií napsat libovolný řádek znaků, který vznikne tak, že za písmenka A , B a C dosadí nějaké formule. Například naše tři axiomy jsou formule a dosazením výroku $1+1=2$ za A , a výroku „Teplota je -273°C .“ za B do axiomu (A1), dostáváme výrok „Jestliže $1+1=2$, potom když je teplota -273°C , tak $1+1=2$.“

Náš kalkul má taky jedno pravidlo, logikové mu říkají *modus ponens*. To říká tohle: Pokud jsi už na výstup napsal formuli A a také formuli $A \Rightarrow B$, můžeš tam napsat také formuli B .

Příklad důkazu v kalkulu*

Ukážeme si, jak pomocí tohoto kalkulu dostaneme formuli $p \Rightarrow p$, která říká „Když je pravda p , tak je pravda p “. Tedy například „Když prší, tak prší,“ což určitě chceme, aby bylo pravda.

Nejdřív použijeme axiom (A2); za B dosadíme $p \Rightarrow p$, a za A a C dosadíme p .

$$(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C)) \quad (\text{A2})$$

$$(p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)) \quad (1)$$

⁷Kdyby Ti zvědavost nedala spát, tady přeci jen jsou i ostatní axiomy:

- (A \wedge 1) $(A \wedge B) \Rightarrow A$
 (A \wedge 1') $(A \wedge B) \Rightarrow B$
 (A \wedge 2) $(A \Rightarrow B) \Rightarrow ((A \Rightarrow C) \Rightarrow (A \Rightarrow (B \wedge C)))$
 (A \vee 1) $A \Rightarrow (A \vee B)$
 (A \vee 1') $B \Rightarrow (A \vee B)$
 (A \vee 2) $(A \Rightarrow C) \Rightarrow ((B \Rightarrow C) \Rightarrow ((A \vee B) \Rightarrow C))$
 (A \Leftrightarrow 1) $((A \Leftrightarrow B) \Rightarrow (A \Rightarrow B))$
 (A \Leftrightarrow 1') $((A \Leftrightarrow B) \Rightarrow (B \Rightarrow A))$
 (A \Leftrightarrow 2) $(A \Rightarrow B) \Rightarrow ((B \Rightarrow A) \Rightarrow A \Leftrightarrow B)$

Potom použijeme axiom (A1) a dosadíme do něj stejně jako v předešlém kroku.

$$A \Rightarrow (B \Rightarrow A) \tag{A1}$$

$$p \Rightarrow ((p \Rightarrow p) \Rightarrow p) \tag{2}$$

Teď můžeme použít pravidlo modus ponens na (1) a (2):

$$\begin{array}{l} \text{(MP)} \quad A \Rightarrow B \quad (p \Rightarrow ((p \Rightarrow p) \Rightarrow p)) \Rightarrow ((p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)) \\ \quad \quad A \quad \quad \quad p \Rightarrow ((p \Rightarrow p) \Rightarrow p) \end{array}$$

$$\frac{B \quad (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p)}{\quad} \tag{3}$$

Teď znovu dosadíme do (A1), ale tentokrát za A i B dosadíme p .

$$A \Rightarrow (B \Rightarrow A) \tag{A1}$$

$$p \Rightarrow (p \Rightarrow p) \tag{4}$$

No a nakonec znovu použijeme modus ponens, tentokrát na (3) a (4)

$$\begin{array}{l} \text{(MP)} \quad A \Rightarrow B \quad (p \Rightarrow (p \Rightarrow p)) \Rightarrow (p \Rightarrow p) \\ \quad \quad A \quad \quad \quad p \Rightarrow (p \Rightarrow p) \end{array}$$

$$\frac{B \quad p \Rightarrow p}{\quad}$$

Vidíme, že dokazování v takovém kalkulu není vůbec jednoduchá věc a rozhodně si dovedeme představit i jednodušší způsoby, jak zjistit, že „Jestliže prší, tak prší.“ je vždycky pravda. Přesto se logikové kalkuly rádi zabývají, dokazují o nich různé věci a vymýšlí další (v některých z nich se dokazuje o něco lépe, ale museli bychom si vysvětlit ještě spoustu dalších neznámých pojmů, abychom se s nimi mohli seznámit). My si bez důkazu uvedeme dvě typické věty o kalkulu výrokové logiky:

Věta. (O korektnosti) *Kalkul výrokové logiky je korektní, to znamená, že každá formule, kterou v něm dokážeme, je skutečně tautologie.*

Věta. (O úplnosti) *Kalkul výrokové logiky je úplný, to znamená, že každou tautologii v něm lze dokázat.*

Cvičení 9. Dokaž větu o korektnosti výrokové logiky, to znamená, ukaž pomocí tabulky, že axiomy jsou tautologie a že pravidlo modus ponens ze dvou tautologií opět odvodí tautologii.

Cvičení 10. Klasická logika má tu zajímavou vlastnost, že když k množině všech tautologií přidáme jedinou formuli, která tautologií není, budeme pomocí dosazení moci odvodit spor⁸ – tedy budeme moci najít dvě formule, které obě vznikly dosazením do nějakých tautologií nebo do naší přidané formule, z nichž jedna je negací druhé.

Nechť je dána formule A, která není tautologií. Řekněte, jaké dvě formule použijete a jaké dosazení na každou z nich provedete, abyste dostali hledaný spor.

Shrnutí*

Dnes jsme se seznámili s klasickou výrokovou logikou. Na jejím příkladu jsme si ukázali, jaký je rozdíl mezi syntaxí, tedy čistě formálním hraní s písmenky, a sémantikou, tedy uvažováním

⁸V matematické řeči se říká, že množina tautologií klasické výrokové logiky je *maximální bezesporná množina*.

o významu. Sémantika je velmi důležitá, protože s její pomocí se mezi sebou můžeme dohodnout, co máme na mysli, když vyslovíme tvrzení určitého tvaru. Viděli jsme, že například u implikace nemusí být úplně snadné takovou dohodu udělat (a ještě se k tomu vrátíme v příštích dílech). Zatímco v oblasti sémantiky výrokové logiky jsme se naučili ověřovat, zda je nějaká formule tautologie či kontradikce (nebo nic z toho) pomocí tabulky, v oblasti syntaxe jsme se seznámili s jedním kalkulem. Abychom nebyli pořád jen úplně klasičtí, předvedli jsme si také tzv. dynamický přístup k logice, ve kterém chápeme ověřování pravdivosti nějakého tvrzení jako hru pro dva hráče.

Další cvičení

Nedej se odradit délkou zadání některých úloh! Není vůbec pravda, že čím delší zadání, tím škaredější úloha ...

Cvícení 11. (podle František Gahér: *Provokácia ako motivácia k štúdiu logiky*) Někdy se stává, že nějakou větu v přirozeném jazyce, tedy v jazyce, kterým se lidé běžně dorozumívají, lze pochopit dvěma různými způsoby. Často je to způsobené tím, že přirozený jazyk nezachycuje dostatečně přesně takzvanou logickou strukturu věty. V této úloze se podíváme na dva takové příklady:

(a) Není-li stanoveno právním předpisem nebo účastníky dohodnuto jinak, jsou podíly všech spoluvlastníků stejné.

(Ak nie je právnym predpisom ustanovené alebo účastníkmi dohodnuté inak, sú podiely všetkých spoluvlastníkov rovnaké.)

Tuto větu lze formalizovat dvěma způsoby: $\neg(A \vee B) \Rightarrow C$, $(\neg A \vee \neg B) \Rightarrow C$.

(b) Není-li možné vadu odstranit a není-li pro ni možné užívat věc dohodnutým způsobem nebo řádně, má nabyvatel právo domáhat se zrušení smlouvy.

(Ak nemožno vadu odstrániť a ak nemožno pre ňu vec užívať dohodnutým spôsobom alebo riadne, je nadobúdateľ oprávnený domáhať sa zrušenia zmluvy.)

Dvě možné formule tentokrát jsou: $(\neg A \wedge \neg(B \vee C)) \Rightarrow D$, $(\neg A \wedge (\neg B \vee \neg C)) \Rightarrow D$.

V obou částech ukaž, že dvě navržené formule nejsou ekvivalentní. (po 1 bodě)

V každém příkladě zkus větu přeformulovat tak, jak by měla znít, aby jednoznačně odpovídala první a druhé navržené formuli. Rozhodni, která z nabízených formulí zachycuje zamýšlený význam věty. (po 1 bodě)

Všimni si, jak byla zamýšlena část věty se strukturou „není-li A nebo B“ („ak nie A alebo B“)! (1 bod)

Cvícení 12. Snem tvůrců logických kalkulek bylo vytvořit systém, ve kterém je každý jednotlivý krok důkazu zcela nezpochybnitelný, zkrátka evidentně správný a není tedy třeba jej zdůvodňovat. Karel Čapek o takovém dokazování napsal:

„O logickém důkazu je jediná pravda, že se nic nedá logicky dokazovat; což vám dokážu logicky. Buď dokazují své tvrzení samými evidentními soudy; ale kdyby mé tvrzení plynulo evidentně z evidentních vět, bylo by samo evidentní, a tu by ovšem naprosto nepotřebovalo být dokazováno. Nebo dokazují své tvrzení větami neevidentními, ale pak bych musel logicky dokazovat všechny tyto věty „usque ad infinitum“, [...] z čehož logicky plyne, že logický důkaz je nemožný; a není-li tento logický důkaz naprosto přesvědčující, vidíte z toho, že logické dokazování opravdu za nice nestojí.“ (Karel Čapek: *Kritika slov*; citováno v Antonín Sochor: *Klasická matematická logika*, str. 21, odkud to cituje Petr Jansa, str. 25)

Odhalte, v čem spočívá klam čili podfuk Čapkovy argumentace!

Cvičení 13. (podle James D. McCawley: Everything that Linguists have Always Wanted to Know about Logic) Jeden vědec z planety Venuše se rozhodl zkoumat, zda různé jazyky mohou fungovat na základě různých logických systémů. Po několika letech studia pozemšťanů (konkrétně svou studii prováděl v Čechách) publikoval článek, ve kterém tvrdil následující:

V jazyce Čechů se zápor vytváří tak, že se před určitý tvar slovesa připojí předpona ne-. Češové přitom často používají věty tvaru „ $A \wedge \neg A$ “, například „Někteří lidé jsou chytrí a někteří lidé nejsou chytrí.“ Z toho je vidět, že v logice Čechů neplatí zákon, který na Venuši zná každé malé dítě, totiž že žádný výrok nemůže být pravdivý současně se svou negací.

Vysvětlete venušskému vědci, v čem se mýlí.

Cvičení 14. (podle James D. McCawley: Everything that Linguists have Always Wanted to Know about Logic)

- Nalezněte formuli obsahující pouze spojky \neg a \Rightarrow , která je ekvivalentní s formulí $(A \wedge B) \Rightarrow A$.
- Rozhodněte, zda je formule $(A \wedge B) \Rightarrow A$ dokazatelná v hilbertovském kalkulu výrokové logiky.
- Rozhodněte, které z následujících vět jsou konjunkcemi a které ne: Plakal a plakal. Samá práce a žádná zábava udělaly z Honzy hrozného škarohlída. Tomáš slíbil Šárce kožešinový kabát a prsten s diamantem. Pět a tři je osm. Pan Novák svezl na valníku Petříka a buďto Anežku nebo Kristýnku.

Návody, nápovědy, řešení a poznámky k některým cvičením

Nápověda ke cvičení 10. Vymyslete, jaké dosazení je třeba udělat, abychom z formule, která není tautologií (tedy není pravdivá při nějakém ohodnocení), dostali kontradikci (formuli nepravdivou při všech ohodnoceních).

Řešení 12. cvičení. „Je nutné zdůraznit, v čem se matematický logik naprosto rozchází s citovaným výrokiem K. Čapka. Parafrazují-li jeho úvahu, mohou také jednoduše „dokázat“, že nikdy nedojdu z Prahy do Prčic: při každém kroku se má pozice změnit jen málo a Prčice jsou od Prahy dost daleko. Chyba Čapkovy úvahy spočívá v tom, že mnoha (byť malými) kroky je možno urazit velkou vzdálenost, což řada účastníků pochodu na uvedenou trasu prakticky prokázala. Analogicky není možno popřít, že i pomocí evidentních důkazových kroků je možno dojít (je-li jich hodně) k tvrzením značně neevidentním.“ (Antonín Sochor: Klasická matematická logika, str. 23; cituje Jansa, str. 25)

Neklasické logiky

Intuicionistická logika a konstruktivismus

„Přestože Brouwer chtěl ukázat, že některé matematické důkazy fungují jinak než logika, lidé si všimli, že Brouwerův argument může ukázat také to, že některé oblasti matematiky sice fungují podle logiky, ale tato logika je odlišná. Někteří dokonce takovou logiku vybudovali a pokusili se předvést, že to je ve skutečnosti logika veškeré matematiky. Nazvali to intuicionistickou logikou.“

(Dan Cryan, Sharron Shatil, Bill Mayblin: Logika, str. 94)

Budování základů matematiky

V druhé polovině 19. století prožívala matematika krizi způsobenou zjištěním, že základy, na kterých se dosud budovalo, nejsou dost pevné a připouštějí různé paradoxy. Gottlob Frege se

domníval, že jediným skutečně pevným a důvěryhodným základem pro matematiku je logika. Vypracoval způsob zápisu a dokazování logických formulí, díky kterému je považován za zakladatele moderní logiky. V té době byla za nejdůležitější část matematiky považována nauka o číslech, aritmetika. Frege se pokusil vybudovat aritmetiku na pevných základech logiky, ale když mělo jeho celoživotní dílo po dvanácti letech práce vyjít tiskem, upozornil ho mladý anglický matematik a filozof Bertrand Russel na závažný problém. V podstatě se jednalo o obdobu známého paradoxu holiče:

Holič ze Sevilly holí právě ty ze sevillských mužů, kteří se neholí sami. Holí se holič sám?

Kdyby se holil, neholil by se, protože holí jen ty, kteří se neholí sami. Ale pokud se neholí sám, musí se holit . . .

Přestože se Fregeho logika nakonec neosvědčila, Russel na Fregeho dílo navázal a stejně jako on se pokusil vybudovat pevné základy pro matematiku. Rozhodl se postupovat stejně jako kdysi dávno Euklidés při budování geometrie: některá zřejmá tvrzení prohlásit za axiomy⁹ a všechna složitější tvrzení dokázat pomocí několika postupů, o jejichž správnosti nemůže být pochyb. V třísvazkovém díle *Principia mathematica* Russel společně s Whiteheadem skutečně touto metodou vybuodoval aritmetiku a teorii čísel. Jaké bylo jejich zděšení, když si uvědomili, že se dopustili téže chyby jako Frege!

Problém, se kterým se Russel potýkal, lze zformulovat takto:

Uvažujme množinu všech množin, které nejsou prvkem sama sebe. Je tato množina svým prvkem?

V této době přichází na scénu německý matematik David Hilbert a prohlašuje: „Současný stav věcí je neudržitelný. Jen si pomyslete, definice a deduktivní metody, kterým se všichni učíme, kterým využíváme a které používáme v matematice, vzor pravdy a jistoty, vedou k absurditám! Pokud je matematické myšlení nedokonalé, kde hledat jistotu a pravdu?“¹⁰

Hilbertovi a jeho následovníkům se pak podařilo vybudovat takové axiomatické systémy, které k absurditám nevedly. Podstatnou roli zde hrálo to, že se vzdali představy, že existuje jen jedna „správná“ matematika. Zhruba řečeno, pro Hilberta je „správná“ každá matematická teorie, která nevede k absurditám. Matematicky řečeno, lze si zvolit zcela libovolné axiomy, na jejichž základě nelze dokázat nějaké tvrzení i s jeho negací. Soubor axiomů, který splňuje toto kritérium, nazveme *bezesporná teorie*, protože jeho důsledkem není žádná formule tvaru $A \wedge \neg A$ (takovou formuli nazýváme *spor*). Podle Hilberta mohou v matematice existovat vzájemně konkurenční teorie a nikomu to nevadí.¹¹

⁹Přestože si matematikové velmi zakládají na tom, aby každé jejich tvrzení bylo odůvodněné, není možné snažit se dokázat skutečně všechna tvrzení. V každém důkaze totiž matematik použije nějaká jednodušší tvrzení, která by opět měl dokázat – a tak až do nekonečna. Abychom se vyhnuli nekonečnému dokazování, musíme se rozhodnout, že některým tvrzením budeme důvěřovat bez důkazu. Taková tvrzení se nazývají *axiomy*; někdy, zejména v souvislosti s řeckou matematikou, se používá též označení *postuláty*.

¹⁰John D. Barrow: *Pí na nebesích*, str. 115.

¹¹Již dříve známým příkladem tohoto počínání bylo zavedení neeuklidovských geometrií. Fyzik by řekl, že nejvýše jedna z těchto teorií je správná, protože nejvýše jedna z nich popisuje svět, ve kterém žijeme. V neeuklidovských geometriích může například být součet úhlů v trojúhelníku i jiný než 180°; fyzikové provádějí velmi přesná měření, aby ověřili, jaký je součet úhlů ve „skutečných trojúhelnících“ (tvořených například třemi hvězdami; přímky jsou přitom vymezeny dráhou světelných paprsků). Matematik se na „skutečné trojúhelníky“ neohlíží: zajímá jej jen to, zda je svět, ve kterém je součet úhlů v trojúhelníku jiný než 180°, vůbec myslitelný, zda by

Důležité je, jak se tyto teorie budují. Začne se tím, že napíšeme několik axiomů. Pak je třeba ukázat, že z těchto axiomů nelze odvodit spor.¹² Při odvozování dalších tvrzení smíme používat pouze určité předem popsané malé krůčky. Příkladem takové teorie je například kalkul výrokové logiky.

Netrvalo dlouho a holandský matematik Luitzen Brouwer vystoupil s tvrzením, že axiomatické budování matematiky je zcela mylné. Jeho argument byl zhruba následující:

Matematika je v první řadě činností matematikovy mysli. Matematik intuitivně ví, s jakými objekty pracuje, třeba co to je přirozené číslo. Jazyk, kterým svoje poznatky vyjadřuje, je jen pomůckou pro dorozumívání. Je zcela mylné vyjít z tohoto jazyka, ten formalizovat a pak se domnívat, že prostým hraním se symboly „provozujeme“ matematiku.¹³ Ovšem přesně toho se dopouštíme, pokud budujeme teorie s axiomy a odvozovacími pravidly: sice se snažíme, aby naše axiomy odrážely naše intuitivní představy o matematických objektech, ale jakmile se jednou rozhodneme pro určitou sadu axiomů a odvozovacích pravidel, můžeme další matematické věty dokazovat zcela mechanicky – tak mechanicky, že by to za nás mohl dělat i počítač. Za axiomy své teorie volíme určité věty, výroky či formule, které se skládají ze slov či symbolů, kterými se snažíme (většinou ne zcela přesně) popsat vlastnosti čísel či jiných matematických objektů; díky tomu místo se samotnými čísly pracujeme s větami o číslech.

Podle Brouwera je třeba matematiku budovat pouze na těch nejzákladnějších matematických představách – *intuicích*, které sdílejí všichni lidé. Pro intuicionisty jsou těmito základními kameny matematiky přirozená čísla a princip matematické indukce, který říká toto: Jestliže číslo 1 má nějakou vlastnost a jestliže platí, že má-li číslo n tuto vlastnost, pak ji má i číslo $n + 1$, tak všechna přirozená čísla mají tuto vlastnost.

Mezi společné intuice všech matematiků patří také základní operace s přirozenými čísly (sčítání, násobení a podobně), které nazýváme *konstrukce*. Matematika je vědou o konstrukcích probíhajících v matematikově mysli. Například tvrzení $1 + 3 = 2 + 2$ pro intuicionisty znamená „provedl jsem v mysli konstrukce odpovídající výrazům $2+2$ a $1+3$ a zjistil jsem, že vedou k témuž výsledku“.

Většina matematiků se domnívá, že nezávisle na nás existuje něco jako svět matematických objektů, který můžeme zkoumat a objevovat. Tento názor na matematiku se často nazývá *matematický platonismus*, protože se podobá Platónově představě, že „někde“ existuje svět dokonalých idejí a předmětů, se kterými se setkáváme v každodenním životě, jsou jen nedokonalými obrazy těchto idejí. Například téměř každé malé dítě ví, že přímka je „nekonečně tenká“ čára, ale žádná přímka nakreslená na papíře tuto vlastnost samozřejmě nemá. Přímky, se kterými se setkáváme, jsou tedy pouze nedokonalými napodobeninami těch „skutečných“ přímek, které „bydlí“ ve světě matematických objektů spolu s čísly, trojúhelníky a množinami.

Podle Brouwera tato představa vedla k používání některých postupů, které nejsou oprávněné. Ze všeho nejvíce se mu nelíbilo dokazování existence matematických objektů metodou nazývanou důkaz sporem. Ten patří k matematickému folklóru už od dob starého Řecka, kdy Eukleidés takto dokázal existenci nekonečného počtu prvočísel:

Předpokládejme, že prvočísel je pouze konečný počet a že $2, 3, 5, \dots, P$ jsou všechna prvočísla.

úvahy o něm nevedly ke sporům a paradoxům.

¹²Ve skutečnosti matematické často tento krok vynechávají a pouze doufají, že jejich teorie jsou bezsporné. Profesor Tomáš Kepka z pražského matfyzu k tomu s oblibou říká: „Věřím, že v matematice je spor, ale je příliš daleko, takže ho nikdy neobjevíme.“

¹³„Matematika je rozumová činnost (jednotlivých) lidí a je jen ne zcela adekvátně sdělitelná jazykem.“ Miroslav Mleziva: *Neklasické logiky*, str. 63.

Pak číslo $Q = (2 \cdot 3 \cdot 5 \cdots P) + 1$ je větší než největší prvočíslo P , takže není prvočíslem, a tedy je dělitelné nějakým prvočíslem. Ale po dělení všemi prvočísly $2, 3, 5, \dots, P$ dává zbytek 1, takže žádným prvočíslem dělitelné není. To je spor – ukázali jsme, že Q je i není dělitelné nějakým prvočíslem. Předpoklad, že prvočísel je jen konečný počet, vede ke sporu, takže prvočísel musí být nekonečný počet.

Brouwerovi se na dokazování sporem nelíbilo, že nedává žádný návod na to, jak najít (zkonstruovat) objekt, jehož existenci dokazujeme. Připomíná to známý vtip o matematikovi a fyzikovi, obou zavřených v kobce s hromadou konzerv. Když se po týdnů žalárník přišel podívat do kobky, kde věznil fyzika, našel ho spokojeného a sytého, protože se mu podařilo konzervy otevřít kouskem ostrého kamene, který uloupl ze stěny. Matematik ležel ve své kobce vyčerpaný a téměř mrtvý hladu; stěny pokryl spoustou výpočtů a nápisů (v kobce byl dostatek bláta, kterým mohl psát) a dole v rohu objevil žalárník dvojité podtrženou větu „Z toho plyne, že KONZERVU LZE OTEVŘÍT!“

Brouwer tvrdil, že matematické objekty neexistují mimo matematikovu mysl. Chceme-li dokázat existenci objektu s určitými vlastnostmi, musíme jej *zkonstruovat* pomocí jednoduchých kroků jako sčítání a násobení. Může se stát, že neumíme zkonstruovat předmět s danou vlastností (a tak dokázat jeho existenci), ani dokázat, že neexistuje. V naší mysli se tedy nemusí nacházet ani konstrukce hledaného předmětu a z ní plynoucí poznání, že existuje, ani poznání, že neexistuje. Vidíme, že o poznatcích v matematikově mysli neplatí *zákon vyloučeného třetího* $A \vee \neg A$, který říká, že každé tvrzení je buďto pravdivé nebo nepravdivé. Jenže matematika je pro Brouwera právě studiem konstrukcí probíhajících v matematikově mysli. Musíme tedy odmítnout klasickou logiku.

Odmítnutí zákona vyloučeného třetího úzce souvisí s intuicionistickým pojetím disjunkce. Intuicionista považuje disjunkci $A \vee B$ za dokázanou pouze v případě, že je dokázán aspoň jeden z výroků A, B . Neintuicionistovi stačí ukázat, že za všech okolností je pravdivý alespoň jeden z výroků A, B – nemusí to ale za všech okolností být ten stejný. Díky tomu lze v klasické logice dokázat $A \vee \neg A$, ač je zřejmé, že nelze zcela obecně dokázat jeden z výroků $A, \neg A$.

Jak už bylo řečeno, považoval Brouwer za neintuitivní také dokazování sporem; tvrdí, že je založeno na zákonu vyloučeného třetího, což jsme viděli na příkladu Euklidova důkazu: tvrzení, že prvočísel je konečný počet, vede ke sporu, takže není pravdivé; v důsledku toho musí být nepravdivé, takže prvočísel není konečný (tj. je nekonečný) počet.

Vidíme, že Brouwer odmítal některé z pilířů klasické logiky; ve skutečnosti odmítal budovat matematiku pomocí jakékoli logiky, protože logika zachycuje pouze vlastnosti našeho jazyka, zatímco matematika by se měla zabývat matematickými intuicemi, které jazykem pouze vyjadřujeme. Někteří Brouwerovi následovníci ale logiku neodmítali tak důrazně jako on. Jedním z nich byl i Arend Heyting, který vytvořil intuicionistickou logiku. Heyting sice tvrdil, že „nikdy nebude možno matematicky přesně dokázat, že daná soustava axiomů skutečně obsahuje všechny platné metody důkazu“¹⁴, ale jeho logika se ujala a patří dnes k nejčastěji studovaným neklasickým logikám.

Kolmogorova logika problémů

Jednou z podivných vlastností intuicionistické logiky je to, že v ní neplatí zákon vyloučeného třetího $A \vee \neg A$. To znamená, že zatímco v klasické logice je pravdivý buďto daný výrok (to když je pravdivý), nebo jeho negace (to když je nepravdivý), v intuicionistické logice existuje ještě

¹⁴ Arend Heyting: *Intuitionism*, Amsterdam 1959, str. 102. Cituje Miroslav Mleziva: *Neklasické logiky*, str. 75.

třetí možnost, totiž že není pravdivý ani A , ani $\neg A$! Výroky intuicionistické logiky tedy nesplňují definici, kterou jsme si uvedli v předchozí kapitole, totiž že výrok je tvrzení, které je vždy buďto pravdivé nebo nepravdivé.

Kolmogorov přišel s myšlenkou neuvažovat o výrocích, ale o problémech. Nadále tedy používal symboly klasické logiky, ale význam symbolu A nebyl „výrok A je pravdivý“, ale „problém A je řešitelný“. Podobně interpretoval i logické spojky:

$\neg A$	problém A je neřešitelný
$A \wedge B$	oba problémy, A i B , jsou řešitelné
$A \vee B$	A je řešitelný nebo B je řešitelný
$A \Rightarrow B$	je-li vyřešen problém A , je vyřešen i problém B
$A \Leftrightarrow B$	zkratka za $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

Co to ale znamená, že nějaký problém je řešitelný? V konstruktivistickém duchu to znamená „umíme zkonstruovat řešení“. Jinak řečeno, řešitelná je pouze ta úloha, ke které je znám postup, jak ji vyřešit. Konstruktivisté nepřipouštějí tvrzení, že daný problém je řešitelný, ale my to ještě neumíme. Neřešitelností problému se pak myslí to, že umíme dokázat, že žádné řešení neexistuje. Zřejmě existuje i třetí možnost: dosud daný problém vyřešit neumíme, ani nevíme, jestli skutečně nějaká řešení má. Zákon vyloučeného třetího $A \vee \neg A$ tedy neplatí.

Podívejme se, jak v této interpretaci dopadne další z důležitých postupů klasické logiky, totiž důkaz sporem. Příslušná formule zní $(\neg A \Rightarrow A) \Rightarrow A$. V klasické logice jsme uvažovali následovně: kdyby bylo pravda $\neg A$, musí podle první implikace být pravda také A , ale A nemůže být současně pravda i nepravda. Je-li tedy implikace $(\neg A \Rightarrow A)$ pravdivá, nemůže být výrok $\neg A$ pravdivý, a podle zákona vyloučeného třetího $A \vee \neg A$ musí být pravdivý výrok A .

Tato formule zachycuje základní myšlenku dokazování sporem: vyjdeme od negace dokazovaného tvrzení a snažíme se ukázat, že z ní plyne nějaký zřejmý nesmysl. V úvaze z předešlého odstavce je tím nesmyslem formule $A \wedge \neg A$; logikové ji obvykle nazývají *spor*. Pokud z $\neg A$ plyne spor, $\neg A$ platit nemůže, a musí tedy platit A .

Implikace $(\neg A \Rightarrow A)$ v Kolmogorově interpretaci říká zhruba toto: umíme-li vyřešit problém $\neg A$, umíme vyřešit i problém A . Řešením problému $\neg A$ je ale důkaz, že problém A je neřešitelný! Kolmogorův výklad formule $(\neg A \Rightarrow A) \Rightarrow A$ tedy je „je-li vyřešena úloha přepracovat jakýkoli důkaz, že problém A nemá řešení, na řešení problému A , umíme nalézt i samotné řešení problému A “. Na první pohled je vidět, že nemáme žádný důvod považovat tuto formuli za platnou¹⁵: řešení úlohy přepracovat důkaz neřešitelnosti A na řešení problému A by nám k řešení problému A pomohlo pouze tehdy, když už bychom měli připravený důkaz její neřešitelnosti.

Do třetice se podívejme na to, jak z pohledu pana Kolmogorova dopadne klasická ekvivalence $A \Leftrightarrow \neg\neg A$. Stejně jako v klasickém případě ekvivalenci považujeme za pouhou zkratku za konjunkci dvou implikací $(A \Rightarrow \neg\neg A) \wedge (\neg\neg A \Rightarrow A)$.

¹⁵O nějaké formuli řekneme, že je *platná*, je-li tautologií.

*Označení *platná formule* se v logice používá v několika významech: 1) jako synonymum pro *pravdivá* (zejména při definování splňování v modelech predikátové logiky se často používá výraz „formule je platná v modelu“ místo „formule je pravdivá při dané interpretaci“); 2) jako synonymum pro *pravdivá nezávisle na ohodnocení proměnných* – to je ve výrokové logice totéž co *tautologická*. (V predikátové logice, která v jistém smyslu obsahuje výrokovou logiku, se pak slůvko *tautologická* používá pro formule, které jsou pravdivé nezávisle na ohodnocení proměnných díky své výrokově-logické struktuře, zatímco *platná* pro formule, u nichž je nutno vzít v potaz jejich predikátově-logickou strukturu a vlastnosti modelu, se kterým právě pracujeme.); 3) ve významu *důsledek teorie* neboli *pravdivá ve všech modelech dané teorie*.

První implikace je platná i při výkladu pomocí řešitelnosti problémů: je-li A „pravdivé“, známe způsob, jak nalézt řešení problému A. Ovšem tím je dokázáno, že není neřešitelný; jinými slovy, je dokázáno, že problém $\neg A$ (tedy úloha dokázat neřešitelnost problému A) je neřešitelný.

Opačná implikace klasicky platí také, ale v Kolmogorově interpretaci ne: $\neg\neg A$ říká, že umíme dokázat, že nelze dokázat neřešitelnost problému A; z toho ještě nijak neplyne, že problém A už umíme vyřešit!¹⁶

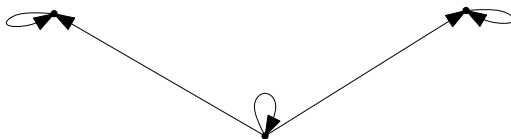
Kripkovské modely intuicionistické logiky¹⁷¹⁸

Odpověď na otázku, zda umíme nějaký problém vyřešit, se samozřejmě může měnit v čase: nejprve nevíme, zda je problém A řešitelný, není tedy pravdivé ani A, ani $\neg A$. Později se třeba dovíme, že není neřešitelný ($\neg\neg A$) a nakonec třeba objevíme řešení (A bude pravdivé). Budeme navíc předpokládat, že jsme dost chytrí na to, abychom nezapomněli nic z toho, co jsme už zjistili.

Uvažujeme tvrzení, která jsou pravdivá nebo nepravdivá (ale my třeba nevíme, co z toho), jejichž pravdivost se nemění v čase; tedy výrok „Prší.“ neuvažujeme, ale výrok „17. prosince roku 1900 v 13.45 nad Národním divadlem v Praze pršelo.“ uvažujeme. Na každý takový výrok se můžeme dívat jako na problém určit, která z možností pravda / nepravda nastává. Způsob, jak zachytit vývoj našich znalostí v čase, se nazývá (*kripkovský*) *model*. Ty časové okamžiky, které nás zajímají, budou reprezentovány takzvanými *možnými světy*.¹⁹ V každém možném světě jsou některá tvrzení pravdivá a některá nepravdivá, ale o některých ještě není rozhodnuto, neplatí zde tedy zákon vyloučeného třetího. Označení *možné světy* se používá proto, že nebudeme uvažovat jen okamžiky odpovídající skutečnému stavu našeho poznání někdy v minulosti nebo v budoucnosti, ale i možnosti, které se neuskutečnily. Podívejme se na takový malíčký model pro výrok „Na Marsu je život.“ Je zvykem kreslit modely tak, že „nejstarší“ možné světy jsou dole. Přibývání času (a tedy i poznatků) se označuje šipkami.

Zjistili jsme, že výrok „Na Marsu je život,“
je nepravdivý.

Zjistili jsme, že výrok „Na Marsu je život,“
je pravdivý.



Dosud nepovažujeme za pravdivý ani
výrok „Na Marsu je život,“ ani výrok
„Na Marsu není život.“

Podobný obrázek bychom si mohli nakreslit pro jakýkoli výrok. To, že v daném možném světě S je o daném tvrzení A známo, že je pravdivé, označíme značkou $S \Vdash A$. To, že dané tvrzení

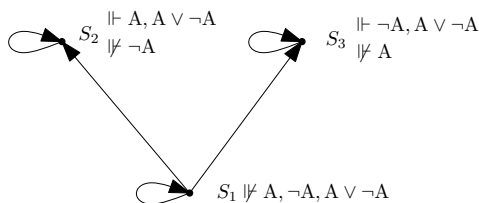
¹⁶Z toho, že víme, že úloha zapsat číslo π s přesností na 10^{100000} desetinných míst není neřešitelná pomocí supervýkonných počítačů budoucnosti, ještě neplyne, že takové počítače máme k dispozici nebo že je umíme sestavit.

¹⁷Místo *kripkovský model* budeme říkat pouze *model*.

¹⁸Tato kapitola se snaží ukázat na příkladech, co to je model a proč vypadá tak, jak vypadá. Záleží na Tobě, zda je pro Tebe snazší přecíst si nejdřív delší a trochu rozvláčná vysvětlení bez velkého množství matematických pojmů v téhle kapitole nebo stručnou a přehlednou, zato zcela abstraktní matematickou definici v kapitole příští.

¹⁹Místo *možný svět* budeme často říkat pouze *svět*.

zatím za pravdivé nepovažujeme, označíme tak, že tuto značku přeškrtneme. Uvědom si, že je rozdíl mezi $S \Vdash \neg A$ a $S \not\vdash A$! Předchozí model tedy bude vypadat takto:



Uvědom si, že (přinejmenším podle intuicionistů) tvrzení tvaru $A \vee B$ můžeme prohlásit za pravdivé pouze tehdy, když už jsme ukázali pravdivost jednoho z výroků A, B . Speciálně v jednom z možných světů v minulém modelu neplatí zákon vyloučeného třetího $A \vee \neg A$!

Tvrzení tvaru $A \wedge B$ můžeme prohlásit za pravdivé, pokud už víme, že obě tvrzení A i B jsou pravdivá. To znamená, že $S \Vdash A \wedge B$ právě, když $S \Vdash A$ a současně $S \Vdash B$.

Důležitým předpokladem tvorby modelů je to, že tvrzení, která jsme jednou prohlásili za pravdivá, budeme za pravdivá považovat už napořád. To znamená, že pokud $S_1 \Vdash A$ a do světa S_2 vede z S_1 šipka (v matematické hatmatilce se říká, že S_2 je z S_1 *dosažitelný*), tak $S_2 \Vdash A$. Matematici této vlastnosti modelů říkají *podmínka perzistence* a je typická právě pro modely intuicionistické logiky. Tedy ještě jednou: o ničem, co jsme jednou prohlásili za pravdivé tvrzení, nemůžeme v budoucnu zjistit, že ve skutečnosti je nepravdivé. Omyly nejsou dovoleny a žádné poznatky se nezapomínají.

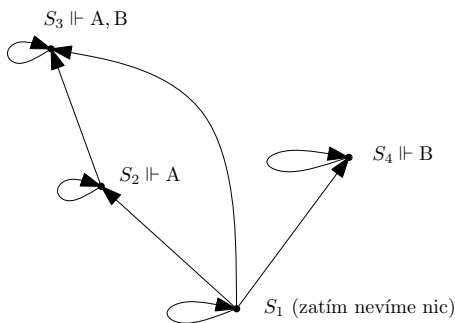
Zajímavější než podmínky pro konjunkci a disjunkci jsou ale podmínky, za kterých můžeme za pravdivou prohlásit nějakou implikaci nebo negaci. Pro účely tvorby modelů začíná budoucnost už v přítomném okamžiku, pokud tedy napíšu „nikdy v budoucnu“, mám na mysli „nyní ani nikdy v budoucnu“. V obrázcích modelů to znázorníme tak, že z každého možného světa vede šipka zpátky do něj samotného, a v matematické řeči řekneme, že relace dosažitelnosti je reflexivní.

Z podmínky perzistence je vidět, že je-li pravdivé $\neg A$, nemůže se stát, že někdy v budoucnu bude pravdivé A . V takovém budoucím okamžiku by totiž bylo pravda $A \wedge \neg A$, což ale v žádném rozumném světě nastat nemůže.

Tuto úvahu lze i obrátit: Negaci $\neg A$ prohlásíme za pravdivou, kdykoli víme, že nikdy v budoucnu neprohlásíme za pravdivou formuli A . To znamená, že pokud v žádném možném světě dosažitelném z S není pravdivé A , můžeme psát $S \Vdash \neg A$.

Intuitivně tušíme, že by implikace $A \Rightarrow B$ měla vypovídat něco o příčinném vztahu mezi A a B . Proto se i podmínka pro pravdivost implikace $A \Rightarrow B$ odvolává na pravdivost v (možných) budoucích okamžicích: pokud v každém budoucím okamžiku bude tato implikace pravdivá klasicky, pak v daném okamžiku je pravdivá i intuicionisticky. Jinak řečeno, pokud z S není dosažitelný žádný možný svět, ve kterém A je pravda a B nikoli (tedy ve kterém je implikace $A \Rightarrow B$ nepravdivá ve smyslu klasické logiky), tak $S \Vdash A \Rightarrow B$.

Tohle vypadá trochu složitě, ale hnedka si to ukážeme na příkladu:



formule	je pravdivá v možných světech	poznámka
A	S_2, S_3	
B	S_3, S_4	
$\neg A$	S_4	ze všech ostatních světů se lze po šípkách dostat někam, kde je pravda A
$\neg B$	nikde!	odevšad se lze po šípkách dostat někam, kde je pravda B
$A \Rightarrow B$	S_3, S_4	z S_1 i S_2 se lze dostat do S_2 , kde je pravda A, ale není pravda B
$\neg\neg A$	S_2, S_3	z S_1 i S_4 se lze dostat do S_4 , kde je pravda $\neg A$
$\neg\neg B$	všude!	odnikud se nelze po šípkách dostat někam, kde je pravda $\neg B$
$A \vee \neg A$	S_2, S_3, S_4	
$B \vee \neg B$	S_3, S_4	S_1 i S_2 jsou protipříklady na zákon vyloučeného třetího
$\neg\neg A \Rightarrow A$	S_1, S_2, S_3, S_4	všude, kde je pravda $\neg\neg A$, je pravda i A
$\neg\neg B \Rightarrow B$	S_3, S_4	z S_1 i S_2 se lze dostat do S_2 , kde je pravda $\neg\neg B$, ale není pravda B

Cvičení 15. Někteří intuicionisté považují formuli $\neg A$ za zkratku za formuli $A \Rightarrow \perp$, kde \perp označuje spor, tedy jakoukoli formuli tvaru $B \wedge \neg B$. Ušetří si tím práci s definováním, kdy je pravdivá formule $\neg A$. Spor nikdy pravdivý není. Ukaž, že takto definovaná negace je pravdivá právě v těch možných světech, ve kterých je pravdivá na základě podmínky popsané v této kapitole.

Cvičení 16. Pravděpodobně Tě to překvapí, ale Brouwerovi nevadilo vyvrácení sporem – pokud nějaké tvrzení vede ke sporu, je jistě nepravdivé, a tedy je pravdivá jeho negace. Znamená to, že Brouwer odmítá formuli $(\neg A \Rightarrow A) \Rightarrow A$ popisující dokazování sporem, ale přijímá formuli $(A \Rightarrow \neg A) \Rightarrow \neg A$. Vysvětli, jak to souvisí s pojetím negace popsaném ve cvičení 15.

V tuto chvíli už tedy tušíš, že model je způsob, jak matematicky zachytit přibývání poznatků a že v něm jsou zachyceny i ty situace, které ve skutečnosti nenastanou, ale třeba by nastat mohly. V následující kapitole si ukážeme, jak se tento typ modelů definuje v řeči matematiky.

* Matematická definice kripkovských modelů²⁰

Relace R na množině W je nějaký vztah, do kterého vstupují vždy dva prvky množiny W .²¹ Například vztah „mít rád“ je relace na množině lidí.²² Dalšími příklady relací jsou vztahy „být otcem“, „být manželem“, „být partnerem“, „být sourozencem“, „chodit do stejné třídy“.

To, že prvky a a b jsou ve vztahu R můžeme značit různými způsoby: $(a, b) \in R$, $R(a, b)$ nebo aRb . Příkladem druhého způsobu zápisu jsou relace „být menší nebo roven než“ a „být větší než“, které znáš ze školy: $1 \leq 3, 3 > 1$. Jiným příkladem téhož zápisu je tvorba jednoduchých vět: maminka má ráda děťátko, paní učitelka má ráda poslušné dítě.

V případě žáků jedné školy je každý žák sám se sebou v relaci „chodit do stejné třídy“; o relaci, která má vlastnost, že každý prvek množiny W (také přezdívané *nosná množina*) je sám se sebou v relaci, řekneme, že je *reflexivní*. Stručně to lze zapsat $\forall x \in W xRx$.

Při tvorbě modelů se budeme nejčastěji setkávat s relací dosažitelnosti, která odpovídá vlastnosti „být později než“.²³ Pro tu platí, že je-li okamžik Q pozdější než okamžik P a ten je zas pozdější než okamžik O , je Q pozdější než O . Matematicky to zapíšeme $(R(O, P) \wedge R(P, Q)) \Rightarrow R(O, Q)$ a řekneme, že R je *tranzitivní relace*.

Kripkovský model se skládá z neprázdné množiny možných světů W s reflexivní a tranzitivní relací dosažitelnosti \leq .²⁴ O každém možném světě S je určeno, které výroky A v něm jsou pravdivé (což značíme $S \models A$) a které pravdivé nejsou (což značíme $S \not\models A$). Toto určení ale není libovolné, musí totiž splňovat následující podmínky:²⁵

- | | | |
|-----------------------------------|-------------------|---|
| • $S_1 \models A, S_1 \leq S_2$ | potom | • $S_2 \models A$ (<i>podmínka perzistence</i>) |
| • $S \models A \wedge B$ | právě tehdy, když | • $S \models A$ a $S \models B$ |
| • $S \models A \vee B$ | právě tehdy, když | • $S \models A$ nebo $S \models B$ |
| • $S_1 \not\models \neg A$ | právě tehdy, když | • $S_1 \leq S_2$ potom $S_2 \not\models A$ |
| • $S_1 \models A \Rightarrow B$ | právě tehdy, když | • $S_1 \leq S_2, S_2 \models A$ potom $S_2 \models B$ |
| • $S \models A \Leftrightarrow B$ | právě tehdy, když | • $S \models A \Rightarrow B$ a $S \models B \Rightarrow A$. |

$S \models A$ čteme „v možném světě S je formule A pravdivá“ nebo „v možném světě S je splněna formule A “. Není-li v možném světě S formule A pravdivá, píšeme $S \not\models A$ a říkáme „v možném světě S není splněna formule A “. Pozor – to není totéž jako „v možném světě S je formule A nepravdivá“ – $S \models \neg A$!

Cvičení 17. Nalezněte model a v něm možný svět S a formuli A takové, že $S \not\models A$, ale přitom není pravda, že $S \models \neg A$.

²⁰Místo *kripkovský model* budeme většinou říkat pouze *model*.

²¹Matematická definice: Relace R na množině W je množina uspořádaných dvojic $R = \{(a_1, b_1), (a_2, b_2), \dots\}$, kde a_1, a_2, \dots a b_1, b_2, \dots jsou prvky množiny W .

²²mít rád = $\{(\text{Pepíček, Mařenka}), (\text{maminka, děťátko}), (\text{paní učitelka, poslušné dítě}), (\text{Mařenka, Honzík}), \dots \}$.

²³Přesněji řečeno, relace dosažitelnosti odpovídá vlastnosti „být totožný s nebo moci nastat později než“.

²⁴Formálně vzato patří k modelu ještě *valua*ce, která každému možnému světu určuje, které atomární formule, tedy výroky, které neobsahují logické spojky, jakési „věty jednoduché“ logiky, v něm jsou pravdivé.

²⁵Vyjmenované podmínky musí být splněny ve **všech** možných světech $S, S_1, S_2 \in W$ a musí platit pro všechny formule A, B !

Budeme-li prvky W kreslit jako tečky v rovině a $S \leq R$ označíme šipkou ze světa S do světa R , znamená reflexivita a tranzitivita relace \leq to, že z každého světa musí vést šipka zpátky do něj a také do všech světů, do kterých se lze dostat po šípkách.

Intuicionistické tautologie a kripkovské protipříklady

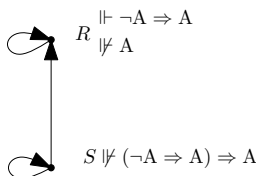
Formuli, která je splněna ve všech možných světech všech modelů, nazveme *intuicionistická tautologie*.

Pokud nějaká formule F není intuicionistickou tautologií, existuje nějaký model a v něm nějaký možný svět, ve kterém dotyčná formule není splněna. Tento model nazveme (*kripkovský protipříklad pro formuli F*).

V předchozích dvou kapitolách jsme si ukázali protipříklady pro formule tvaru $A \vee \neg A$ a také pro formule tvaru $\neg\neg A \Rightarrow A$ (pouze jsme ji napsali jako $\neg\neg B \Rightarrow B$). Ukažme si ještě na příkladu formule $(\neg A \Rightarrow A) \Rightarrow A$ ²⁶, jak se protipříklady vytváří. Vždycky můžeme začít tím, že si nakreslíme možný svět, ve kterém zadaná formule není splněna:

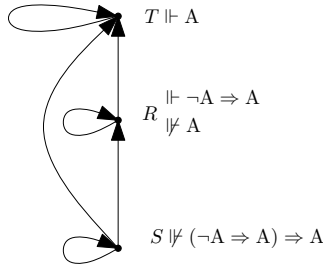
$$S \not\models (\neg A \Rightarrow A) \Rightarrow A$$

Podle definice modelu ovšem implikace neplatí v možném světě S jen tehdy, je-li z něj dosažitelný nějaký svět R , ve kterém je její první člen pravdivý a její druhý člen není pravdivý (tedy kde tato implikace neplatí ani klasicky).



Vypadá to nadějně: zdá se, že z R není (po šípkách) dosažitelný žádný možný svět, kde je splněno $\neg A$, takže implikace $\neg A \Rightarrow A$ je splněna triviálně. Zbývá ověřit, že všechny ostatní podmínky z definice modelu jsou splněny. . . Jenže pozor! Z R není dosažitelný ani žádný svět, kde je splněno A , takže v R je splněno $\neg A$! Přeci jen je z R dosažitelný nějaký svět, kde $\neg A$ je pravdivé, takže musíme zkontrolovat pravdivost implikace $\neg A \Rightarrow A$. Podle té by ve světě R muselo být pravdivé také A , a to by znamenalo, že tam je pravda A a $\neg A$ současně. To ale není možné vzhledem k podmínce pro spojku \neg z definice modelu. Znamená to snad, že svět R do našeho obrázku nemůžeme přikreslit? Ne; problémů se zbavíme tak, že přikreslíme svět T dosažitelný z R , kde bude splněno A (tím v R přestane být splněno $\neg A$ a tudíž budeme moci použít úvahu z předminulého odstavce):

²⁶To je formule popisující dokazování sporem, o které byla řeč již v kapitole o Kolmogorově interpretaci formulí jako výroků o řešitelnosti problémů.



Ve čtvrté seriálové úloze budeš mít za úkol ověřit, že se skutečně jedná o model, a tedy o hledaný protipříklad.

Cvičení 18. Sestroj protipříklad k následující formuli:

$$((\neg\neg A \Rightarrow A) \Rightarrow (A \vee \neg A)) \vee (\neg\neg A \Rightarrow A).$$

Hilbertovský kalkul intuicionistické logiky

První práce, ve které Brouwer popisuje intuicionistickou filozofii, pochází z roku 1907, ovšem širší veřejnost s ní začal seznamovat až po roce 1912, kdy se stal profesorem. Ve 20. letech dospěl Hilbert k závěru, že to Brouwer opravdu myslí vážně, a začal intuicionistickou filozofii považovat za hrozbu pro budoucí vývoj matematiky. Právem upozorňoval na to, že intuicionisté se snaží „ustavit matematiku tak, že shodí přes palubu vše, co se jim nehodí, a uvalí na to blokádu. Cílem je rozložit naši vědeckou disciplínu a riskovat ztrátu našich nejcennějších objevů.“²⁷ Přeci jen, intuicionistická matematika je ochuzena o některé z nejdůležitějších důkazových prostředků používaných od starověku.

V té době se pracovalo s tabulkami pro spojky klasické logiky, ale po prvotních neúspěšných pokusech vytvořit vícehodnotové tabulky pro intuicionistickou logiku Kurt Gödel roku 1932 ukázal, že neexistuje konečněhodnotová tabulka, ve které by tautologiemi byly právě ty formule, které konstruktivisté uznávali. Kripkovská sémantika vznikla až v 60. letech.

Ve snaze formálněji zachytit a popsat, které matematické postupy jsou pro intuicionisty přípustné a které ne, byli tedy logikové zpočátku odkázáni na logické kalkuly. To, jak se takové kalkuly tvoří a jak se s nimi pracuje, důkladně ukázali Russel a Whitehead ve svém díle *Principia Mathematica*.

Ukážeme si jeden z mnoha hilbertovských kalkulů²⁸ pro intuicionistickou logiku. Axiomy lze zvolit mnoha různými způsoby – zde si ukážeme způsob zajímavý tím, že přidáme-li k těmto axiomům ještě jeden jediný další axiom, dostaneme hilbertovský kalkul pro klasickou logiku (ovšem s trochu jinými axiomami, než které jsme vyjmenovali posledně).

Při formulaci axiomů pro klasickou logiku jsme vycházeli z toho, že každou formuli lze ekvivalentně zapsat pouze pomocí implikace a negace. Stačilo tedy napsat axiomy obsahující tyto dvě spojky. Kdybychom chtěli dokázat nějakou formuli obsahující i jiné spojky, našli bychom nejprve ekvivalentní formuli pouze se spojkami \Rightarrow a \neg a tu pak dokázali v kalkulu.

²⁷ John D. Barrow: *Pí na nebesích*, str. 212.

²⁸ Označení *hilbertovský kalkul* se používá pro kalkuly, ve kterých se pracuje s jednotlivými formulami. Tyto kalkuly obvykle obsahují pravidlo modus ponens a pravidlo o dosazování. Existují i jiné typy kalkulů.

V intuicionistické logice něco takového udělat nemůžeme, protože nelze vybrat několik spojek, pomocí kterých by šlo ekvivalentně napsat všechny formule. Říkáme, že spojky $\wedge, \vee, \Rightarrow$ a \neg jsou v intuicionistické logice *nezávislé*. Pouze spojce \Leftrightarrow rozumíme jako zkratce za konjunkci dvou implikací a proto si nebudeme uvádět axiomy, které ji obsahují.

Axiomy obsahující implikaci a negaci jsou čtyři:

- (Ai1) $A \Rightarrow (B \Rightarrow A)$
- (Ai2) $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$
- (Ai3) $(B \Rightarrow \neg A) \Rightarrow (A \Rightarrow \neg B)$
- (Ai4) $A \Rightarrow (\neg A \Rightarrow B)$

První dva z těchto axiomů jsou totožné s axiomy (A1) a (A2) z kalkulu klasické výrokové logiky. Ta obsahovala ještě axiom

$$(A3) (\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B),$$

který jsme nahradili podobně vypadajícím axiomem (Ai3). Jenže se ukazuje, že s pomocí axiomu (Ai3) lze dokázat mnohem méně formulí – dokonce ani formuli (Ai4), kterou bychom s pomocí axiomu (A3) snadno dokázali.

Slibovaný jediný axiom, který bychom museli přidat k axiomům (Ai1), ... (Ai4), abychom mohli odvodit všechny tautologie klasické logiky, také obsahuje pouze spojky \Rightarrow a \neg :

$$(A^*) (\neg A \Rightarrow B) \Rightarrow ((\neg A \Rightarrow \neg B) \Rightarrow A)$$

To je formule pro *důkaz sporem*, jen napsaná o něco obecněji, než jsme to udělali v předchozích částech této kapitoly: když za B dosadíme A, dostaneme

$$(\heartsuit) (\neg A \Rightarrow A) \Rightarrow ((\neg A \Rightarrow \neg A) \Rightarrow A).$$

Ovšem $(\neg A \Rightarrow \neg A)$ je v kalkulu intuicionistické logiky dokazatelná – vznikne totiž substitucí $\neg A$ za p do $p \Rightarrow p$, o které jsme si v minulé kapitole ukázali, že je dokazatelná pouze pomocí axiomů (A1) a (A2) – tedy (Ai1) a (Ai2). Díky tomu formule (\heartsuit) implikuje formuli $(\neg A \Rightarrow A) \Rightarrow A$, kterou jsme nazvali *důkaz sporem*.

Cvičení 19. Ukažte podrobně, jak lze pomocí axiomů (Ai1), ... (Ai4), formule $(\neg A \Rightarrow A) \Rightarrow ((\neg A \Rightarrow \neg A) \Rightarrow A)$ a pravidla modus ponens dokázat formuli $(\neg A \Rightarrow A) \Rightarrow A$.

První nápověda: Pro zpřehlednění důkazu si zaveď následující označení: $a = (\neg A \Rightarrow A)$, $b = (\neg A \Rightarrow \neg A)$, $c = A$.

Druhá nápověda: Jako mezikrok odvoď formuli $(a \Rightarrow b) \Rightarrow (a \Rightarrow c)$.

Axiomy pro ostatní spojky se shodují s příslušnými axiomy v kalkulu klasické logiky:

- (A \wedge 1) $(A \wedge B) \Rightarrow A$
- (A \wedge 1') $(A \wedge B) \Rightarrow B$
- (A \wedge 2) $(A \Rightarrow B) \Rightarrow ((A \Rightarrow C) \Rightarrow (A \Rightarrow (B \wedge C)))$
- (A \vee 1) $A \Rightarrow (A \vee B)$
- (A \vee 1') $B \Rightarrow (A \vee B)$
- (A \vee 2) $(A \Rightarrow C) \Rightarrow ((B \Rightarrow C) \Rightarrow ((A \vee B) \Rightarrow C))$
- (A \Leftrightarrow 1) $((A \Leftrightarrow B) \Rightarrow (A \Rightarrow B))$
- (A \Leftrightarrow 1') $((A \Leftrightarrow B) \Rightarrow (B \Rightarrow A))$
- (A \Leftrightarrow 2) $(A \Rightarrow B) \Rightarrow ((B \Rightarrow A) \Rightarrow A \Leftrightarrow B)$

Stejně jako v klasické logice do kalkulu patří také pravidlo modus ponens:

Pokud jsi už na výstup napsal formuli A a také formuli $A \Rightarrow B$, můžeš tam napsat také formuli B.

Také pro intuicionistickou logiku platí věta o korektnosti a věta o úplnosti:

Věta 20. (O korektnosti a o úplnosti) *Kalkul intuicionistické logiky je korektní a úplný vzhledem ke kripkovským modelům intuicionistické logiky, to znamená, že v něm lze dokázat všechny intuicionistické tautologie (úplnost) a nic víc (korektnost).*

Cvičení 21.²⁹ Rozhodněte, zda následující formule jsou intuicionisticky tautologické. Pokud ne, sestrojte protipříklad. Pokud ano, dokažte je v kalkulu intuicionistické logiky nebo jinak zdůvodněte, že protipříklad sestrojiti nelze. Uvědomte si (ukažte tabulkou), že všechny tyto formule jsou z hlediska klasické logiky tautologiemi.

- $\neg A \vee \neg B \vee (A \Rightarrow B) \vee (B \Rightarrow A)$
- $\neg\neg\neg\neg B \Rightarrow B$
- $(\neg\neg A \Rightarrow A) \Rightarrow (A \vee \neg A)$
- $(A \vee \neg A) \Rightarrow (\neg\neg A \Rightarrow A)$

Důležitost důkazu sporem v matematice

Z předchozího výkladu by se mohlo zdát, že matematikové by důkaz sporem měli považovat za jakousi berličku, kterou je radno používat jen v nejnnutnějších případech (pokud vůbec). Ve skutečnosti se názory různí – jen velmi málo matematiků důkaz sporem odmítá úplně, ale mnozí skutečně dávají přednost konstruktivním důkazům existence (tedy důkazům, v nichž je nalezen objekt s požadovanými vlastnostmi) před nekonstruktivními (tedy důkazy sporem). Najdou se i matematikové, kteří dokazování sporem obdivují a považují je za velmi elegantní. Citujme alespoň G. H. Hardyho, který krásu matematického důkazu přirovnává ke krásě dobře zahrané šachové partie: „Důkaz je proveden metodou *reductio ad absurdum*³⁰ a *reductio ad absurdum*, které Eukleidés tak miloval, je jednou z nejlepších zbraní matematika. Je to mnohem vybranější gambit než jakýkoli gambit šachový: šachista může jako oběť nabídnout pěšce či figuru, ale matematik nabízí *hru*.“³¹

Neklasické logiky

Epistemické logiky

V minulém dílu seriálu jsme se seznámili s intuicionistickou logikou. Viděli jsme, že ji lze považovat za logiku, která popisuje, jak se v čase vyvíjí naše (pravdivé) poznání světa. Intuicionismus ovšem vznikl v rámci matematiky a také dnes je významný zejména jako popis určitého přístupu k matematickým důkazům; pro popis znalostí se častěji používají různé logiky, které označujeme souhrnným názvem *epistemické*³².

Cvičení v tomto dílu jsou velmi důležitá, proto Ti vřele doporučuji se na ně podívat; ke všem je na konci uvedeno řešení.

Pexeso a základní pojmy epistemické logiky

Představte si dva hráče, kteří právě rozložili na stůl šedesát čtyři kartičky pexesa a chystají se hrát. Jaké jsou v tuto chvíli jejich znalosti o rozložení kartiček?

²⁹Vítězslav Švejdar: Cvičení ke kurzu Logika II, část III, cvičení 3.

³⁰*Reductio ad absurdum* je latinský název pro důkaz sporem: dovedení k nesmyslnému.

³¹G. H. Hardy: Obrana matematikova, str. 86.

³²Řecké slovo *epistémé* znamená poznání. *Epistemologie* je část filozofie, která se zabývá procesem poznávání a vztahem poznání ke skutečnosti. Typická otázka zní: „Jak víme, že je naše poznání pravdivé?“

Na první pohled by se mohlo zdát, že žádné: všechny kartičky jsou obráceny rubem vzhůru. Ale ve skutečnosti toho hráči vědí poměrně hodně – vědí, že kartičky skrývají třicet dva různých obrázků a že každý obrázek se vyskytuje právě dvakrát. Navíc jsou si oba vědomi toho, že druhý hráč také ví, že kartičky skrývají třicet dva párů obrázků. Můžeme říci, že mají nějakou *sdílenou znalost*.³³

Obvykle se hraje tak, že hráč, kterému se podaří otočit dvě stejné kartičky, může pokračovat otočením další dvojice. Aby byla hra zajímavější, dohodli se naši hráči na tom, že podaří-li se jednomu otočit dvě dvojice kartiček po sobě, nechá hrát druhého hráče.

Představme si nyní situaci, která může nastat uprostřed hry. Hráč, který je právě na tahu, zná polohu tří dvojic stejných obrázků (jsou to autíčko, beruška a citron). Musí si z nich vybrat dvě a doufat, že tu třetí mu protihráč nevyfoukne ve svém tahu. Protihráč si nepamatuje, kde se nachází jeden z obrázků citronu, ale pamatuje si dobře, kde najde autíčka a berušky; jenže to náš hráč neví, takže otočí dvojice autíček a citronů. V této situaci sice oba hráči věděli, kde najít obrázky berušky, ale první hráč nevěděl, že to druhý hráč také ví – proto nehovoříme o společné znalosti, ale pouze o *znalosti všech*.³⁴ Všimni si, že kdyby se jednalo o sdílenou znalost, první hráč by se rozhodl lépe.

Nakonec se podíváme na znalosti hráčů v okamžiku, kdy na hrací ploše zbývá jediná kartička, kterou zatím nikdo neotočil (tj. všechny ostatní už někdy otočené byly). Pokud si hráč pamatuje, co je na všech ostatních kartičkách, nejspíš si domyslí také to, co se nachází na této zbývající kartičce. Ale co když se náhodou stane, že na hrací ploše je ještě velký počet kartiček? Pokud hráč není opravdu bravurní, snadno se mu přihodí, že si ani nevšimne, že už zná všechny kartičky kromě jediné. Prostě otočí některou z dvojic, které zná, a nebude se snažit si vzpomenout na to, co je na ostatních kartičkách. Vidíme tedy, že ačkoli si hráč poslední obrázek možná domyslí, rozhodně to není jisté.

Zkusme si nyní rozmyslet, s jakými překážkami se budeme muset vypořádat při tvorbě logického systému, který by nám pomohl analyzovat znalosti hráčů v průběhu hry pexeso:

(1) Jsou zde dva hráči, jejichž znalosti se mohou lišit. V epistemických logikách se obvykle hovoří o různých *agentech* – těmi mohou být agenti tajných služeb, kteří sbírají informace všeho druhu, stejně jako roboti, kteří jsou naprogramováni k tomu, aby prováděli meteorologická měření a sestavovali z nich předpověď počasí. Důležitou součástí epistemických logik je právě studium *multiagentních systémů*.

Agenti si za určitých okolností mohou předávat informace, takže znalost jednoho se rázem stane i znalostí druhého. Některé znalosti mají všichni agenti (*znalost všech*) a o některých dokonce všichni vědí, že je mají všichni ostatní a že všichni ostatní také vědí, že je mají všichni ostatní a tak dále (*sdílená znalost*).

(2) Hráči získávají znalosti nejen přímo (otočením kartiček), ale mohou si je také domýšlet. Ve většině her i skutečných životních situacích se vyplatí předpokládat, že protihráč si skutečně domyslel všechny informace, které mohl; proto se v mnoha epistemických logikách nepracuje přímo se znalostmi agentů, ale s takzvanými *implicitními znalostmi* – s poznatky, které si agenti mohou domyslet. Mluvíme o *logické vševědoucnosti agentů*: předpokládáme, že znají všechny logické důsledky svých znalostí.

V praxi se ovšem s vševědoucími agenty setkáváme jen zřídkka. Kdyby například hráči šachu byli vševědoucí, dokázali by si představit všechny budoucí tahy hry a nikdy by se jim nestalo, že přehlédnou soupeřovu šanci připravit je o královnu! Dokonce ani počítačové programy nedokáží

³³V angličtině se používá označení *common knowledge* či *mutual knowledge*.

³⁴V angličtině se používá vyjádření *everyone knows that*.

v rozumném čase prozkoumat všechny možnosti, jak by se hra mohla dál vyvíjet, a vybrat z nich tu nejlepší.

(3) Znalosti hráčů se vyvíjejí v čase – hráči otáčejí nové kartičky a zjišťují, co na nich je zakresleno, a také zapomínají, co bylo na kartičkách otočených dříve. Logické vlastnosti znalostí vyvíjejících se v čase a znalostí o minulosti a budoucnosti se snaží zachytit *epistemicko-temporální logika*.

Výhody sdílených znalostí

Na první pohled se znalost všech liší od sdílené znalosti jen maličko, ale už jsme viděli, že za určitých okolností může tento rozdíl hrát velkou roli.

Představ si, že jsi s kamarádem na Matějské pouti a najednou se jeden druhému ztratíte. Chvilku bloudíš mezi atrakcemi a houfy lidí, ale svého kamaráda nikde nevidíš. Jako na potvoru u sebe zrovna nemáš mobil, takže to vypadá, že už se nejspíš nikdy nevidíte. Co v tuhle chvíli uděláš?

Je dost možné, že budeš přemýšlet asi takhle: „Petr se chtěl jít podívat do Domu hrůzy. Možná bych ho tedy měl(a) jít hledat tam . . . Ale ne, určitě už si všiml, že jsme se ztratili, a teď mě hledá. Kde by mě asi Petr mohl hledat? Říkal(a) jsem mu, že se mi líbí Obří kolo. Nejspíš mě tedy hledá tam . . . Ale Petrovi už určitě došlo, že ho také hledám, a že tedy nejsem u Obřího kola. Protože ví, že ho hledám, určitě půjde tam, kde si myslí, že bych ho mohla hledat. Naposledy jsme se viděli u horské dráhy, takže si nejspíš řekl, že ho budu hledat tam. Když se lidé jeden druhému ztratí, mají se vrátit tam, kde se viděli naposledy . . .“ Jenže u horské dráhy je zrovna neproniknutelný houf lidí, fronta se točí kolem vstupu do atrakce a je jasné, že tady se jen tak nenajdete. „Dobrá,“ řekneš si, „když jsme spolu na Matějské, vždycky si jdeme dát cukrovou vatu, a tady zrovna vidím nápis *CUKROVÁ VATA 100 m*. Jestli mě Petr hledal tady, určitě ho taky napadlo, že tam bychom se našli nejsnáze.“ Díky tomu, že sdílíte znalost, že si obvykle dáváte cukrovou vatu, se během pár minut potkáte u zmiňovaného stánku. Důležitou roli zde hraje nejen to, že víš, že oba máte rádi cukrovou vatu, ale také to, že Ti je jasné, že to ví i Petr, a také to, že Petr ví, že ty víš, že on to ví . . . Zkrátka a dobře, je to vaše sdílená znalost a tak se můžete spolehnout na to, že oba budete jednat na jejím základě.

V následující situaci ze života číšník záměrně přetvoří znalost obou ve sdílenou znalost:

Nešikovný číšník zakopne a vylije omáčku na bílou saténovou blůzku jedné dámy. Dáma na něj vyděšená a rozzlobená zírá a číšník omluvně vykoktá: „Promiňte, byla to moje chyba.“

Co vlastně této dámě říká? Z jejího výrazu mu musí být jasné, že návštěvnice moc dobře ví, že to byla jeho chyba. Ve skutečnosti číšník dává najevo, že ON ví, že to byla jeho chyba. Tím, že to řekne nahlas, se ujišťuje, že ona ví, že on ví, že to byla jeho chyba. V tuto chvíli se znalost obou (totiž že to byla jeho chyba) stává jejich sdílenou znalostí (oba vědí, že ten druhý ví, že oni vědí atd.).³⁵

V životě jakékoli společnosti hrají sdílené znalosti zcela zásadní roli. Pokud se občas stane, že se sdílení znalostí naruší, mívá to katastrofální následky, jak ukazuje následující příklad:

Chodec přechází ulici na přechodu bez semaforu. Ví, že má přednost, takže vstoupí do vozovky, ačkoli se blíží auto. Řidič je bohužel cizinec a neví, že v naší zemi mají chodci na přechodu

³⁵Sdílená znalost je taková, ve které řetězec „ona ví, že on ví, že . . .“ může pokračovat libovolně daleko. Důležitou roli zde hraje i to, že oba vědí, že ten druhý si z už řečeného umí domyslet i to, že vědí, že on ví, že . . . Například číšník si domýšlí, že dáma po jeho prohlášení ví, že on ví, že to byla jeho chyba. Dáma si zase domýšlí, že on ví, že ona si domyslela, že on ví, že to byla jeho chyba.

přednost; vjede do křižovatky a jen tak tak stihne stočit automobil stranou a narazí do sloupu. Vyděšený chodec se ptá: „Copak jste nevěděl, že budu přecházet?!“; předpokládal totiž, že znalost předpisů je jejich sdílenou znalostí, tedy že řidič ví, že chodec bude jednat na základě svého práva přejít křižovatku jako první.

Všimni si, že kdyby si chodec nemyslel, že znalost předpisů je jejich sdílenou znalostí, ale že je jen znalostí jich obou, možná by do vozovky nevstoupil. Například pokud se chodec pozdě v noci vrací domů a slyší blížící se hřmění motoru silného sportovního auta, bude raději předpokládat, že si ho (nejspíš opilý) řidič nevšimne.

Syntaxe epistemické logiky

Nejjednodušším způsobem, jak označit tvrzení „Agent A zná fakt F .“ je symbol $K_A F$.³⁶ V epistemické logice jsou fakty reprezentovány výroky, takže $K_A F$ čteme také jako „Agent A ví, že výrok F je pravdivý.“, zatímco $K_A \neg F$ znamená „ A ví, že výrok F je nepravdivý.“

Pokud si agenty očísľujeme A_1, \dots, A_n , budeme místo $K_{A_1} F$ psát pouze $K_1 F$. Znalost všech budeme značit EF , takže můžeme napsat

$$EF \Leftrightarrow K_1 F \wedge K_2 F \wedge \dots \wedge K_n F.$$

Sdílenou znalost všech agentů budeme značit CF .³⁷

Cvičení 22. Zkus napsat formuli obsahující symboly K a E ekvivalentní s formulí CF . Pak zkus tuto formuli upravit tak, aby neobsahovala symbol E .

Zatímco ve skutečném životě se může snadno stát, že se domnívám, že něco vím, ale ve skutečnosti se mýlím, v logice tuto možnost obvykle nepřipouštíme. Věděním se tedy rozumí *pravdivá domněnka*. To, že se agent A domnívá, že fakt F je pravdivý, značíme $B_A F$.³⁸ Domněnka může být pravdivá nebo nepravdivá.

Kripkovské modely znalostí jediného agenta

Představme si pro začátek situaci, ve které vystupuje jediný agent, třeba agent 007. Jeho znalosti jsou omezené. V důsledku toho nedokáže o všech větech rozhodnout, zda jsou pravdivé nebo nepravdivé. Například ví, že nebezpečný náklad se nachází ve vlaku z Moskvy do Petrohradu, ale neví, zda je pravda, že kapitán Flint zakopal svůj poklad na Ostrově pokladů. Svět, ve kterém se poklad nachází na Ostrově pokladů, je pro něj *možný*, ale rozhodně není *nutný*. V zásadě si dokáže představit dva možné stavy světa podle toho, zda poklad je či není na ostrově. V logice se různé stavy světa označují pojmem *možné světy*: každý možný svět reprezentuje jednu z alternativ, jak náš svět může vypadat.

My všichni „bydlíme“ v jednom z *možných světů*, nikdo z nás ale není schopen přesně určit, ve kterém! Nanejvýš se můžeme omezit na množinu možných světů, které se shodují s našimi znalostmi. Tím je určena *relace dosažitelnosti*: z našeho světa jsou dosažitelné všechny, které odpovídají našim znalostem a domněnkám.³⁹ Například pro agenta 007 jsou dosažitelné všechny možné světy, v nichž se nebezpečný náklad nachází ve vlaku z Moskvy do Petrohradu, a světy, v nichž je tento náklad v Londýně, pro něj dosažitelné nejsou.

³⁶Písmenko K pochází z anglického *knowledge*.

³⁷Písmenko E pochází z anglického *everyone knows that*, C pochází z *common knowledge*.

³⁸Písmenko B pochází z anglického *belief*.

³⁹A žádné světy, které našim znalostem a domněnkám neodpovídají, dosažitelné nejsou.

Aby naše povídání o možných světech nesklouzlo do nekonečných debat o tom, které světy jsou možné a které možné nejsou, povězme si hned ze začátku, že z hlediska logiky není možný svět ničím více a ničím méně než objektem, o kterém víme, které výroky v něm jsou pravdivé a které další možné světy z něj jsou dosažitelné. Přitom předpokládáme, že pravdivost složitějších výroků závisí na pravdivosti jednodušších výroků stejně jako v klasické logice. Například svět, ve kterém je pravda „Mars je planeta.“ a také je pravda „Eustach je planeta.“ a přitom není pravda „Mars a Eustach jsou planety.“, nebudeme považovat za možný.⁴⁰

Definice 23. *Určitou množinu možných světů spolu s pevně danou relací dosažitelnosti nazýváme Dkripkovský model nebo prostě Dmodel.*

Příklad 24. Omezme se na následující dvojici výroků:

(V) „Lady Smithová je v hotelu Interkontinental.“

(W) „Agent 007 si vždycky ví rady.“

Vzhledem k těmto výroky rozlišujeme celkem čtyři možné světy:

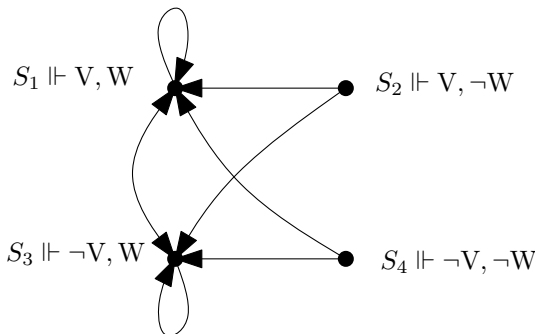
(S₁) $V \wedge W$

(S₂) $V \wedge \neg W$

(S₃) $\neg V \wedge W$

(S₄) $\neg V \wedge \neg W$

Agent 007 si prvním výrokem není moc jist, ale zato se domnívá, že druhý výrok je pravdivý: $B_{007}W$. Ať už se nachází v kterémkoli ze světů S_1, \dots, S_4 , jsou pro něj dosažitelné právě světy s lichým číslem. Pokud označíme relaci dosažitelnosti šipkou mezi dvěma možnými světy, bude obrázek vypadat takto:

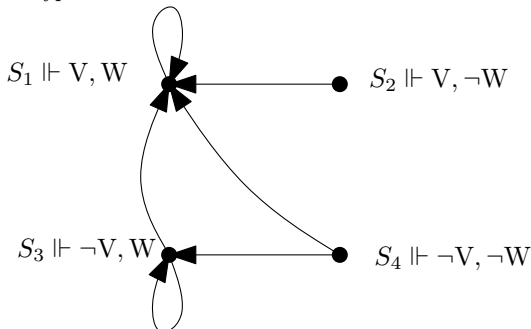


Všimni si, že náš model neříká nic o tom, ve kterém možném světě se agent vlastně nachází. Je-li jeho domněnka správná, pak je to určitě v jednom ze světů S_1 a S_3 – ale z hlediska modelu se klidně může stát, že jeho domněnky pravdivé nejsou. Šipka z daného možného světa S tedy říká, které stavy světa považuje agent za možné, pokud se nachází ve světě S .

Příklad 25. V předchozím příkladě svět, ve kterém se agent nacházel, nehrál žádnou roli. Přijmeme pro jednoduchost předpoklad, že pokud je lady Smithová v hotelu Interkontinental,

⁴⁰Díky tomu, že složitější výroky musí splňovat pravidla klasické logiky, stačí pro popis možného světa určit pravdivostní hodnoty atomických výroků, tedy těch výroků, které nelze rozdělit na jednodušší výroky spojené pomocí logických spojek.

tak si toho agent všimne, tedy $V \Rightarrow K_{007}V$. Nadále předpokládejme také $B_{007}W$.⁴¹ Relace dosažitelnosti teď bude vypadat takto:



Vidíme, že v tomto případě už to, které světy budou dosažitelné, závisí také na tom, ve kterém světě se agent nachází (z některých světů je dosažitelný jen svět S_1 , z některých navíc také svět S_3). Všimni si také, že *čím více má agent poznatků, tím méně světů je dosažitelných*.

Cvičení 26. Urči, jak by vypadala relace dosažitelnosti, kdyby platilo $V \Leftrightarrow B_{007}V$ – tedy nejen, že agent ví, že lady je v hotelu, pokud tomu tak skutečně je, ale pokud si agent myslí, že lady je v hotelu, tak se nemylí. Stále předpokládáme, že $B_{007}W$.

Explicitní a implicitní znalosti

Zatím jsme popisovali, které možné světy budeme na základě svých znalostí považovat za dosažitelné. Položme si nyní opačnou otázku: je dáno, které možné světy jsou dosažitelné; o kterých výrocích řekneme, že mohou být pravdivé? A o kterých řekneme, že jsou nutně pravdivé?

Odpověď není složitá: „A může být pravdivé“ je pravdivé tehdy, je-li dosažitelný alespoň jeden možný svět, kde je A pravdivé; „nutně A“ je pravdivé tehdy, je-li A pravdivé ve všech dosažitelných možných světech. V logice obvykle možnost značíme $\diamond A$, nutnost $\square A$.

Agent tedy považuje za „možná pravdivé“ ty výroky, které jsou pravdivé v některém možném světě, ve kterém jsou pravdivé všechny jeho poznatky; takové výroky jsou *konzistentní*⁴² s jeho poznatkem. Za nutně pravdivé považuje ty výroky, které jsou pravdivé ve všech možných světech, ve kterých jsou pravdivé všechny jeho poznatky; takové výroky totiž z jeho poznatků vyplývají.⁴³

Cvičení 27. Rozhodni, jak vypadá relace dosažitelnosti (vzhledem k našim znalostem) mezi následujícími možnými světy:

- S Víme, že existuje alespoň 6 planet; ve skutečnosti existuje právě 9 planet (ale to nevíme).
- T Víme, že existuje alespoň 6 planet; ve skutečnosti existuje právě 8 planet (ale to nevíme).
- U Víme, že existuje právě 9 planet.⁴⁴

⁴¹Agent si ovšem není jist, že je-li lady Smithová v hotelu, tak by si toho všiml, takže pokud tam není, neví o její přítomnosti či nepřítomnosti nic.

⁴²V logice se místo *konzistentní* používá často české slovo *slučitelné*.

⁴³Připomeňme si definici *vyplývání*: výrok A vyplývá z množiny výroků M , pokud není možné, aby všechny výroky z M byly pravdivé, ale výrok A byl nepravdivý.

⁴⁴Ve světě U existuje právě 9 planet a ve světě V existuje právě 17 planet, protože věděním rozumíme pravdivé poznání.

V Víme, že existuje právě 17 planet.

Pomocí pravdivosti v dosažitelných možných světech určí, ve kterých z těchto světů jsou pravdivé výroky „Nutně existuje 9 planet.“ a „Nutně existuje 8 planet.“ (Předpokládej, že neexistují žádné možné světy kromě S, T, U, V .) Ověř, že $\Box W$ je pravdivé právě tehdy, když víme W .

V předchozím cvičení jsme viděli, že $KW \Leftrightarrow \Box W$. Implikace $KW \Rightarrow \Box W$ odpovídá tomu, že dosažitelné jsou ty světy, které jsou v souladu s našimi poznatky. Pokud tedy agent ví, že platí výrok V (KW), musí být V pravdivé ve všech dosažitelných světech ($\Box W$). Opačná implikace zdaleka tam samozřejmě není: může existovat nějaký výrok, který je pravdivý ve všech možných světech dosažitelných z daného světa, ale agent si toho nemusí vůbec být vědom. Takovým výrokiem je třeba výrok „Je pravda, že jestliže svítí sluníčko, tak prší, nebo je pravda, že jestliže prší, tak svítí sluníčko.“ Jako tautologie je tento výrok pravdivý ve všech možných světech (nejen v těch dosažitelných).

Přestože implikace $\Box W \Rightarrow KW$ tedy pro skutečné agenty neplatí, logici se rozhodli budovat logiku tak, jako by platila. Za chvíli uvidíme, jaké to má výhody, teď se podívejme na to, jak se tím změní význam symbolu $K_A W$.

$\Box W$ je pravda v těch světech, z nichž jsou dosažitelné pouze světy, v nichž je V pravdivý. $\Box W$ tedy znamená „Výrok V vyplývá ze znalostí agenta A .“ Je-li agent dost chytrý, dokáže si výrok V domyslet – V patří k jeho *implicitním znalostem*, to znamená k těm, které by mohl mít.

Vzpomeň si na příklad šachistů: samozřejmě je možné, že můj protihráč si nevšiml, že mi bude moci sebrat věž, pokud s ní neuhnu, ale bude chytrější, když budu předpokládat, že to ví. V této situaci uvažuji o implicitních znalostech svého protihráče.

Od této chvíle bude symbol KV označovat implicitní znalosti agenta.

Vlastnosti kripkovských modelů

Cvičení 28. Urči, jak by vypadala relace dosažitelnosti, pokud by agent oplýval dokonalými poznávacími schopnostmi, tedy pokud by pro každý výrok V platilo $V \Rightarrow KV$. (Můžeš zkusit uvažovat obecný případ, ve kterém je každý svět popsán pomocí n atomárních výroků.) Předpokládej, že agent nemá žádné nepravdivé domněnky.

Cvičení 29. Jak se změní výsledek předešlého cvičení, pokud budeme předpokládat pouze $V \Rightarrow BV$ a připustíme, že agent má i nepravdivé domněnky?

Jak vypadá model znalostí nějakého agenta? Protože pro každý výrok V platí $KV \Rightarrow V$ (za znalost považujeme jen pravdivé výroky), je vždycky skutečný stav světa konzistentní s agentovými poznatky. V obrázku se to projeví tak, že každý svět je dosažitelný sám ze sebe.

Formule $KV \Rightarrow V$ bývá nazývána *axiom T* a logikové ji považují za formuli, která zachycuje rozdíl mezi vědomostmi a pouhými domněnkami.

Dejme tomu, že agent ví, že pokud svítí sluníčko a prší, bude duha ($V \Rightarrow W$) a navíc si všimne, že právě svítí sluníčko a prší (V). Za těchto okolností nás nepřekvapí, že agent ví, že bude duha (W). Tuto schopnost agenta domyslet si na základě svých znalostí něco, co z nich logicky vyplývá, zachycuje formule $(K(V \Rightarrow W) \wedge KV) \Rightarrow KW$, nazývaná *axiom K* nebo také *axiom logické racionality agenta*.

Formality

Podívejme se nyní v rychlosti na to, jak vypadají kalkuly epistemické logiky.

Axiomy

(KL) Všechny tautologie klasické logiky⁴⁵.

(K) Axiom logické racionality $(K(V \Rightarrow W) \wedge KV) \Rightarrow KW$.

Pravidla

(MP) Z dokázaných formulí V , $V \Rightarrow W$ odvoď formuli W .

modus ponens

(Nec) Pokud jsi už dokázal formuli V , odvoď formuli KV .

necesitace

Další axiomy pro různé logiky⁴⁶

Následující axiomy bývají někdy přidávány k základnímu axiomu K , když chceme zachytit další důležité vlastnosti některých agentů a jejich znalostí:⁴⁷

(T) $KV \Rightarrow V$

vědomosti jsou pravdivé

(D) $\neg B(V \wedge \neg V)$

agent nemá rozporné domněnky

(4) $KV \Rightarrow KKV$

pozitivní introspekce: vím, co vím

(5) $\neg KV \Rightarrow K\neg KV$

negativní introspekce: vím, co nevím

Axiom T bývá považován za základní axiom *epistemické logiky*, protože formalizuje definici vědění, která se používá již od středověku: *vědění je pravdivá domněnka*.⁴⁸

Axiom D je považován za základní axiom *doxastické logiky*, protože popisuje důležitou vlastnost většiny našich domněnek a názorů: nezastáváme současně opačné názory, nevěříme současně výroku i jeho negaci. V tomto smyslu jsou epistemické a doxastické logiky dvouhodnotové: danému výroku buďto věřím, nebo nevěřím, ale nemůžu mu (trochu) věřit a (trochu) nevěřit.

Axiomy 4 a 5 zachycují důležité vlastnosti lidských agentů: o každém faktu jsme schopni rozhodnout, zda patří k našim znalostem nebo ne.⁴⁹

Cvičení 30. Ukaž, že jestliže K splňuje axiom T, tak splňuje i obměněnou verzi axiomu D: $\neg K(V \wedge \neg V)$!

Připomeňme si, že relace R na množině W je nějaký vztah, do kterého vstupují vždy dva prvky množiny W . Například vztah „ A je bratr B “ je relace na množině lidí: Jeník je bratrem Petra, Petr je bratrem Jeníka, Jeník je bratrem Mařenky (ale Mařenka není bratrem Jeníka!).⁵⁰

Definice 31.⁵¹ *DKripkovský model (pro jednoho agenta) je tvořen Dmnožinou možných světů M a Drelací dosažitelnosti \leq na množině M .*

⁴⁵Stačilo by přijmout za axiomy všechny axiomy klasické logiky.

⁴⁶Názvy T, D, 4 a 5 jsou tradiční názvy jednotlivých axiomů.

⁴⁷Pozor: ne všichni agenti mají tyto vlastnosti (viz např. axiom **D** a cvičení 9)!

⁴⁸Tuto definici vědění přesněji zachycuje definice $KA \Leftrightarrow (BA \wedge A)$. Někteří myslitelé, například Occam, kladli na vědění silnější požadavky: *vědění je odůvodněná pravdivá domněnka*, což by zachytila formule $KA \Leftrightarrow (BA \wedge A \wedge JA)$, kde JA označuje tvrzení „Agent má důvod věřit, že A .“

⁴⁹Neplatí to ovšem pokaždé: jak často se mi už stalo, že jsem si nechala vyprávět nějaký vtíp, a po vyslovení pointy jsem se chytla za hlavu: „Jéé, vždyť já ten vtíp vlastně už znám!“

⁵⁰Matematikové považují relaci za množinu uspořádaných dvojic, což umožňuje předchodí tři věty napsat stručněji: „být bratrem“ = $\{(Jeník, Petr), (Petr, Jeník), (Jeník, Mařenka), \dots\}$.

⁵¹Tato definice se podobá definici kripkovských modelů intuicionistické logiky, ale je jednodušší: na relaci dosažitelnosti neklade žádné zvláštní podmínky a podmínky pro pravdivost $\neg V$, $V \wedge W$, $V \vee W$, $V \Rightarrow W$ a $V \Leftrightarrow W$ jsou stejné jako v klasické logice. Slovo „kripkovský“ budeme často pro přehlednost vynechávat; v tomto textu jiné než kripkovské modely neuvažujeme.

O každém možném světě $S \in M$ je určeno, které výroky V v něm jsou pravdivé (což značíme $S \models V$) a které pravdivé nejsou (což značíme $S \not\models V$). Toto určení ale není libovolné, musí totiž splňovat následující podmínky:⁵²

- $S \models \neg V$ právě tehdy, když neplatí $S \models V$
- $S \models V \wedge W$ právě tehdy, když $S \models V$ a $S \models W$
- $S \models V \vee W$ právě tehdy, když $S \models V$ nebo $S \models W$
- $S \models V \Rightarrow W$ právě tehdy, když $S \models W$ nebo neplatí $S \models V$
- $S \models V \Leftrightarrow W$ právě tehdy, když $S \models V \Rightarrow W$ a $S \models W \Rightarrow V$
- $S \models KV$ právě tehdy, když $S \leq T$ potom $T \models V$

- (i) $S \models V$ čteme „v možném světě S je výrok V pravdivý“ nebo „v možném světě S je splněna formule V “.
- (ii) $S \not\models V$ čteme „v možném světě S není splněna formule V “ nebo „v možném světě S není výrok V pravdivý“.
- (iii) KV čteme „agent ví, že V je pravdivé“ nebo „agent si může domyslet V “.

Uvedli jsme si axiomy epistemické logiky (nebo lépe řečeno epistemických logik, protože vybereme-li si jen některé z těchto axiomů, dostaneme různé epistemické logiky, které budou popisovat různé typy agentů) a definici kripkovských modelů. K čemu nám ale axiomy a modely jsou?

Logikové se pokoušejí zachytit důležité vlastnosti určitých slov a pojmů, které běžně používáme. V epistemické logice jsou takovými důležitými pojmy „vědět“ či „znát“ a „věřit“. Tyto pojmy mají určitý význam, který je znám všem lidem, kteří umí mluvit příslušným jazykem; jako u většiny slov, ani význam pojmů vědět a věřit není přesně ohraničený, což nám nijak nebrání je používat v naší každodenní řeči i ve filozofii.

Jedním z problémů, na které logikové narážejí ve své snaze zachytit vlastnosti pojmů, jsou právě drobné nuance jejich významů. Logikové proto vytvářejí umělé, formální systémy, které tento význam více či méně přesně zachycují a zpřesňují. Prvním krokem na jejich cestě bývá volba vhodných symbolů – v našem případě to je symbol K pro pojem „ví, že“, symbol B pro „věří“ a symboly $\neg, \wedge, \vee, \Rightarrow$ a \Leftrightarrow pro logické spojky. Symboly ovšem nezachycují význam pojmů, které symbolizují; jen umožňují zapsat věty přehlednějším a stručnějším způsobem a vynechat z nich vše, co není důležité pro tu část významu, která nás v dané chvíli zajímá.

Druhým krokem je volba vhodného formálního systému, který by nějakým způsobem zachytil, jak se zvolené pojmy chovají; obvykle říkáme, že zvolený systém *formalizuje logické vlastnosti* daných pojmů. V zásadě existují dvě cesty, jak to udělat:

- (1) *syntaktická cesta* Můžeme se pokusit vybrat axiomy, které by popsaly všechny důležité vlastnosti vybraných pojmů. Například axiom $T: KV \Rightarrow V$ zachycuje důležitou vlastnost, že vědění je pravdivé poznání. Když budeme mít štěstí, najdeme takovou skupinu axiomů, která bude popisovat všechny důležité vlastnosti zvoleného pojmu.
- (2) *sémantická cesta* Pokusíme se nějak zachytit, za jakých okolností budeme nějakou větu s daným pojmem považovat za pravdivou a za jakých okolností za nepravdivou. Opět máme k dispozici několik způsobů, jak to udělat:

⁵²Vyjmenované podmínky musí být splněny ve *všech* možných světech $S, T \in M$ a musí platit pro všechny výroky $V, W!$

(2a) *tabulková metoda* V případě logických spojek nám výborně posloužily tabulky popisující pravdivostní hodnotu složeného výroku v závislosti na pravdivostních hodnotách jednodušších výroků.

(2b) *kripkovské modely* Význam zvolených pojmů lze často dobře interpretovat v řeči možných světů: například to, že něco vím, znamená, že umím vyloučit určité možnosti, jak by mohl vypadat náš svět, a naopak určité možnosti považují za pravděpodobné.

V případě epistemické logiky máme tedy k dispozici axiomy a kripkovské modely. Obojí se snaží nějak zachytit vlastnosti pojmů „vím“ a „věřím“. Je jasné, že výsledek našeho snažení, totiž volba axiomů a definice kripkovského modelu, se už příliš nepodobá tomu, s čím jsme začali, totiž větám v běžné lidské řeči a jejich významu. Přiznejme si, že žádnému smrtelníkovi kromě logiků neříká věta „Vím, že sluníčko je žluté.“ totéž jako věta „Považuji za přípustné (dosažitelné) pouze ty možné světy, v nichž je sluníčko žluté.“ Proto je pro logiky důležitá otázka, zda aspoň axiomy a modely popisují tytéž vlastnosti vybraných pojmů.

Podobně jako v deváté seriálové úloze můžeme také pro obě pravidla ukázat, že s jejich pomocí dokážeme pouze formule, které platí ve všech modelech. Výsledkem je následující důležitá věta:

Věta 32. (o korektnosti) *Každá formule, kterou lze dokázat s použitím tautologií klasické logiky, axiomu K a pravidel modus ponens a necesitace, je pravdivá ve všech světech všech kripkovských modelů.*

Obrácením předchozí věty vznikne *věta o úplnosti*, která říká, že lze dokázat všechny formule, které jsou pravdivé ve všech světech všech modelů.

* Rozdíly mezi intuicionistickou a epistemickou logikou

V tuto chvíli máš nejspíš pěkný zmatek v tom, jaký je vlastně rozdíl mezi intuicionistickou logikou a logikou epistemickou. Obě přeci zachycují významné vlastnosti znalostí, v obou je řeč o kripkovských modelech a možných světech ... Podívejme se tedy podrobněji na různé rozdíly mezi těmito dvěma logikami.

1. Rozdíly v jazyce

Jazykem nějaké logiky rozumíme množinu symbolů, které se v té logice používají. V intuicionistické logice je jazyk tvořen logickými symboly $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$, pomocnými symboly $()$ a proměnnými pro výroky A, B, \dots V epistemické logice se vyskytuje navíc symbol K , v doxastické⁵³ logice symbol B .

2. Rozdíly v zachycení času

Intuicionistická logika věrně zachycuje, jak naše znalosti přibývají v čase: připouští okamžiky, ve kterých nemáme žádné znalosti a říká, že jednou nabyté pravdivé znalosti se neztrácejí (podmínka perzistence).

Naproti tomu epistemická logika popisuje znalosti agenta v jednom časovém okamžiku. Umožňuje určit, jaké důsledky by agent mohl odvodit ze svých aktuálních znalostí a které stavy světa považuje za možné a které za nemožné.

Kdybychom chtěli použít epistemickou logiku k popisu toho, jak se znalosti agenta vyvíjejí v čase, museli bychom do ní přidat další symboly, například FA („Někdy v budoucnosti bude pravdivý výrok A .“) a PA („Někdy v minulosti byl pravdivý výrok A .“) Pomocí těchto symbolů bychom mohli vyjádřit tvrzení jako $FKPA$ („Agent bude vědět, že přišlo.“).

3. Rozdíly v definici kripkovských modelů

⁵³Doxastické logiky jsou ty, které popisují naše domněnky.

V důsledku toho, že kripkovské modely intuicionistické logiky mají zachycovat přibývání agentových znalostí v průběhu času, musí splňovat podmínku perzistence: $S \Vdash V$ a zároveň $S \leq T$ implikuje $T \Vdash V$ – žádný výrok, o kterém jsme poznali, že je pravdivý, už pravdivým být nepřestane. (Je třeba mít na paměti, že intuicionistická logika popisuje znalosti, které nejsou závislé na čase. Výrok „příši“ může samozřejmě být někdy pravdivý a někdy ne, ale tímto typem výroků se intuicionisté nezabývají.)

V důsledku podmínky perzistence nelze $S \Vdash \neg V$ definovat jednoduše jako „není pravda, že $S \Vdash V$ “⁵⁴ a $S \Vdash V \Rightarrow W$ definovat jednoduše jako „není pravda, že $V \wedge \neg W$ “.

Naopak v epistemické logice se definice pravdivosti formulí tvaru $\neg V$ a $V \Rightarrow W$ v daném možném světě neodvolává na žádné jiné možné světy.

4. Rozdíly v množině tautologií

Z předchozího bodu vyplývá, že v epistemické logice se implikace a negace chovají úplně stejně jako v klasické logice, zatímco intuicionistická logika je vůči těmto spojkám „přísnější“ (klade si více podmínek, aby uznala nějakou implikaci či negaci za pravdivou). Také z formulace axiomů vidíme, že všechno, co je pravda v intuicionistické logice, je pravda i v klasické (ale ne naopak), a všechno, co je pravda v klasické, je pravda i v epistemické (což platí i obráceně pro formule, které neobsahují nově přidané symboly K a B). Takže zatímco intuicionistická logika je obsažena v klasické a má méně tautologií, epistemická klasickou obsahuje a tautologií má více.

Modifikace epistemických logik pro multiagentní systémy

Zatím jsme popsali, jak vypadají modely znalostí jediného agenta a jak vypadají kalkuly, které tyto znalosti popisují. Nyní se pokusíme epistemickou logiku rozšířit tak, aby umožňovala popsat znalosti více agentů. Skupinám více agentů se v informatice a logice říká *multiagentní systémy*; může to být třeba skupina agentů tajné služby, skupina robotů, kteří mají společně vyplnit nějaký úkol (například různé moduly výzkumné sondy na Marsu) nebo skupina mravenečků, kteří společně staví mraveniště. Typické pro tyto systémy je to, že znalosti jednotlivých agentů se mohou lišit, ale za určitých okolností si mohou znalosti i předávat.

Již jsme řekli, že formálně zavedeme pro každého agenta A_i jeho vlastní symbol K_i , takže $K_i V$ bude znamenat, že agent A_i zná (pravdivý) výrok V . Pomocí těchto symbolů můžeme zachytit i složité agentovy úvahy o tom, co všechno vědí ostatní agenti: například $K_1 K_2 K_1$ *svítí sluníčko* znamená že agent A_1 si může říct „ A_2 ví, že vím, že svítí sluníčko.“

Definice 33. *DKripkovský model pro množinu agentů A_1, \dots, A_n je tvořen množinou možných světů M a množinou relací dosažitelnosti \leq_1, \dots, \leq_n na množině M . K modelu patří ještě valuační, která každému možnému světu určuje, které atomární formule, tedy výroky, které neobsahují logické spojky, jakési „věty jednoduché“ logiky, v něm jsou pravdivé. Pravdivost složitějších*

⁵⁴Může se totiž stát, že pro nějaký svět S není pravda, že $S \Vdash V$, ale je z něj dosažitelný svět T , ve kterém je pravda, že $T \Vdash V$. Z toho je vidět, že navrhovaná definice negace by narušila podmínku perzistence.

výroků musí splňovat následující podmínky:⁵⁵

- | | | |
|----------------------------------|-------------------|---|
| • $S \Vdash \neg V$ | právě tehdy, když | neplatí $S \Vdash V$ |
| • $S \Vdash V \wedge W$ | právě tehdy, když | $S \Vdash V$ a $S \Vdash W$ |
| • $S \Vdash V \vee W$ | právě tehdy, když | $S \Vdash V$ nebo $S \Vdash W$ |
| • $S \Vdash V \Rightarrow W$ | právě tehdy, když | $S \Vdash W$ nebo neplatí $S \Vdash V$ |
| • $S \Vdash V \Leftrightarrow W$ | právě tehdy, když | $S \Vdash V \Rightarrow W$ a $S \Vdash W \Rightarrow V$ |
| • $S \Vdash K_i V$ | právě tehdy, když | $S \leq {}_i T$ potom $T \Vdash V$ |

Cvičení 34. Představte si sondu, která přistála na Marsu. Po povrchu planety se nyní pohybuje drobný robot, který má za úkol provést základní chemický rozbor místní atmosféry a vzorků hornin. Robot všechny výsledky měření okamžitě odesílá mateřské sondě a ta je opět odesílá vědcům na Zem (v případě, že by se robot nešťastnou náhodou zranil, zapadl do nějaké díry v zemi či se prostě rozbil, budou mít vědci k dispozici alespoň ty informace, které již stihl odeslat). Formálně toto předávání informací zachycují implikace $K_R \Rightarrow K_S$ a $K_S \Rightarrow K_V$.⁵⁶ Jak se tyto implikace projeví v příslušném kripkovském modelu?

Směsice poznámek a citátů

Sdílené znalosti ztracených v davu

Jedním z prvních, kdo se začali zabývat významem sdílených znalostí pro naše sociální interakce, byl Thomas Schelling v díle Strategie konfliktu z roku 1960. Zde čteme:

„Ztratí-li muž svou ženu v nákupním středisku, aniž by se předtím domluvili, kde se potkají, pokud se rozdělí, je velká naděje, že se opět brzy najdou. Pravděpodobně si každý z nich vzpomene na nějaké místo, kde by se evidentně mohli potkat, tak evidentně, že si oba budou jisti, že to je „evidentní“ pro ně pro oba. Žádný z nich se nebude snažit prostě uhodnout, kam ten druhý půjde, což by bylo tam, kam by hádal, že ten druhý bude hádat, že (ten první) půjde, a tak donekonečna. Nejenom *Co bych dělal na jejím místě?* ale *Co bych dělal, kdybych na jejím místě přemýšlel, co by udělala, kdyby přemýšlela, co bych udělal já na jejím místě...*“

Thomas Schelling: The Strategy of Conflict

V roce 1969 Lewis popsal tento problém v jazyce teorie her. Jedná se o strategickou hru, v níž se hráči rozhodují současně. Představte si, že obchodní dům, ve kterém se nachází manželé Petr a Pavlína Novákoví, má čtyři patra. V každém poličku následující tabulky je zapsáno, kolik uspokojení získá Petr (1. položka) a kolik Pavlína (2. položka), jestliže Petr bude v patře odpovídajícím danému řádku a Pavlína v patře odpovídajícím danému sloupečku.

Petr \ Pavlína	1. patro	2. patro	3. patro	4. patro
1. patro	(1,1)	(0,0)	(0,0)	(0,0)
2. patro	(0,0)	(1,1)	(0,0)	(0,0)
3. patro	(0,0)	(0,0)	(1,1)	(0,0)
4. patro	(0,0)	(0,0)	(0,0)	(1,1)

⁵⁵Vyjmenované podmínky musí být splněny ve *všech* možných světech $S, T \in M$ a musí platit pro všechny formule V, W a pro všechny agenty A_1, \dots, A_n !

⁵⁶Robot, Sonda, Vědci. Zanedbáváme zde skutečnost, že předání informací trvá nějaký čas.

Vidíme, že pro oba „hráče“ této jednoduché hry jsou výhodné tytéž situace – totiž ty, v nichž se sejdou ve stejném patře. Tyto situace se nazývají *Nashovy rovnovážné pozice*, protože pokud se jich podaří dosáhnout, žádný z hráčů už nebude mít důvod danou situaci nějak „vylepšovat“.⁵⁷

Je jasné, že pro Petra náročnost této hry spočívá ve výběru patra, do kterého patra půjde. Za jakých okolností se rozhodne jít do 2. patra? Bude to tehdy, když bude mít přiměřenou jistotu, že do téhož patra půjde i Pavlína (jistota 1. řádu). Takovou přiměřenou jistotu může mít jen tehdy, když bude mít přiměřenou jistotu, že Pavlína má přiměřenou jistotu, že on půjde do 2. patra. K tomu, aby Pavlína měla přiměřenou jistotu, že Petr půjde do 2. patra (jistota 2. řádu), potřebuje mít přiměřenou jistotu, že Petr má přiměřenou jistotu, že ona půjde do 2. patra (jistota 3. řádu), takže celkově Petr potřebuje mít jistotu 4. řádu, *žedots* Kdyby tak Petr věděl, že Pavlína ví, že on ví, že ona ví, že on ví, *dots* že půjde do 2. patra, bylo by to dobrým důvodem tam jít.

Přesně tímto způsobem Lewis definuje *sdílenou znalost* Z skupiny agentů. Sdílená znalost je taková znalost, že pro každé n a pro každou posloupnost agentů A_1, A_2, \dots, A_n (tentýž agent se v posloupnosti může opakovat vícekrát), platí: „ A_1 ví, že A_2 ví, že ... že A_n ví, že Z .“ Speciálně pro $n = 1$ dostáváme, že všichni agenti vědí, že Z .

Hra dvou vězňů

Z předešlého vyprávění by se mohlo zdát, že Nashovy rovnovážné pozice vždy odpovídají těm nejvýhodnějším celkovým výsledkům hry. Ukažme si hru, ve které Nashova rovnovážná pozice je pro oba hráče nevýhodná – a přesto v ní hra skončí, budou-li hráči hrát rozumně!

Ty a tvůj parťák jste obviněni z krádeže. Každý z vás je uvězněn v jiné cele, takže se nemůžete nijak domlouvat. Tvůj právník ti sdělil, že pokud se ani jeden z vás nepřizná, budete oba na základě částečných důkazů odsouzeni ke dvěma letům vězení. Pokud se přiznáš a budeš svědčit proti svému parťákovvi, zatímco on se nepřizná, dostaneš pouze roční trest, zatímco tvůj parťák bude odsouzen na čtyři roky vězení. Pokud se přiznáte oba, odsedíte si v chládku tři roky. Víš také, že tvůj parťák je ve stejné situaci jako ty.

Celou situaci můžeš zachytit tabulkou:

já \ parťák	přizná se	nepřizná se
přiznám se	(3, 3)	(1, 4)
nepřiznám se	(4, 1)	(2, 2)

Pravděpodobně budeš uvažovat asi takto: „Bez ohledu na to, jestli se můj parťák přizná nebo ne, je pro mě výhodnější se přiznat – v obou případech dostanu o dva roky nižší trest.“

Protože tvůj parťák uvažuje úplně stejně, oba se přiznáte a odsedíte si tři roky za mřížemi. Při tom by pro vás pro oba bylo výhodnější, kdyby se ani jeden z vás nepřiznal (odseděli byste si pouze dva roky)! Jenže pozice (2,2) není rovnovážná – alespoň jeden z vás by si mohl polepšit tím, že by se přiznal. Proto tato situace pravděpodobně nenastane – kdyby to vypadalo, že se nechceš přiznat, tvůj parťák zpozoruje svou příležitost, přizná se a bude svědčit proti tobě. Tobě nezbude, než se přiznat také (přeci jen je lepší odsedět si tři roky než čtyři) ...

Cvičení 35. Zkus najít situace ze skutečného života, v nichž Nashova rovnovážná pozice (tedy pozice, ve které si už nikdo ze zúčastněných nemůže nijak polepšit) je pro všechny méně výhodná než nějaká jiná pozice, ve které si ale někdo může polepšit (nejspíš na úkor ostatních).

⁵⁷ Anglický výraz je *Nash equilibrium*.

Nápovědy, řešení a poznámky k některým cvičením

Řešení cvičení 22

Průslušná formule by musela vypadat asi takto: $CF \Leftrightarrow EF \wedge EEF \wedge EEEF \wedge \dots$. Všechny výskyty symbolu E můžeme zkusit nahradit podle definice:

$$\begin{aligned}
 CF &\Leftrightarrow K_1F \wedge K_2F \wedge \dots \wedge K_nF \wedge \\
 &\quad K_1K_1F \wedge K_1K_2F \wedge \dots \wedge K_1K_nF \wedge K_2K_1F \wedge K_2K_2F \wedge \dots \wedge K_nK_nF \wedge \\
 &\quad K_1K_1K_1F \wedge \dots \wedge K_1K_nK_1F \wedge K_1K_nK_2F \wedge \dots \wedge K_nK_nK_nF \wedge \\
 &\quad \dots
 \end{aligned}$$

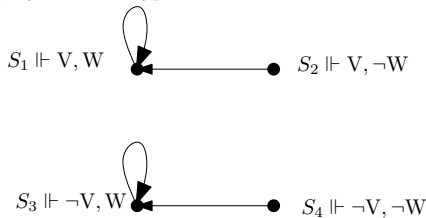
Poznámka ke cvičení 22

Obě formule, které jsme napsali výše, jsou nekonečné. V logice není zvykem pracovat s nekonečnými formulami a navíc je otázkou, jestli skuteční agenti, kteří jsou koneční, mohou podle této definice někdy sdílet znalosti s jinými agenty. Pokud totiž agenti A_1 a A_2 sdílí znalost F , měl by každý z nich mít nekonečné množství znalostí tvaru $K_1F, K_1K_2F, K_1K_2K_1F, \dots$! Při tom je zřejmé, že například lidé mají s jinými lidmi mnoho společných vlastností i přes to, že jejich mozek je schopen ukládat pouze konečné množství informace.

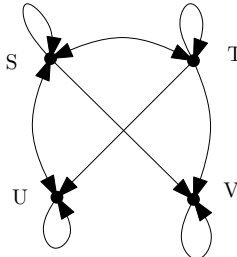
Podobně jako v jiných problémech, se kterými se v epistemické logice setkáme, i zde je řešením prohlásit, že agenti nemusí mít všechny tyto znalosti aktuálně „uložené v paměti“, ale musí je mít potenciálně či implicitně – na základě svých znalostí a schopností by si je mohli kdykoli odvodit či domyslet.

Řešení cvičení 26

V případě, že $V \Leftrightarrow B_{007}V$, by obrázek vypadal takto:



Řešení cvičení 27



Řešení cvičení 28

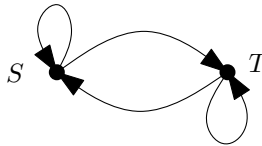
Vědění je vždy pravdivá znalost, tedy $KV \Rightarrow V$. Všimni si, že *pokud připouštíme pouze pravdivé znalosti, tak je z každého světa dosažitelný alespoň on sám*.⁵⁸

Předpokládejme, že každé dva možné světy se liší pravdivostí alespoň jednoho výroku. Pokud je agent vševědoucí a nikdy se nemýlí ($V \Leftrightarrow KV$), je z každého světa dosažitelný pouze tento svět sám. V případě se čtyřmi světy vypadá obrázek takto:



Všimni si, že v tomto případě je agent schopen přesně určit, ve kterém možném světě se nachází – jen jediný svět je pro něj dosažitelný.

Kdybychom připustili možnost, že ve dvou možných světech S a T jsou pravdivé přesně tytéž výroky, vypadal by obrázek takto:



Řešení cvičení 29

Pokud připustíme, že agent má všechny pravdivé poznatky ($V \Rightarrow BV$) a také některé nepravdivé domněnky (nastane případ $\neg V \wedge BV$), nebude z daného možného světa dosažitelný ani on sám, ani žádný jiný. Ve chvíli, kdy agent věří nějakému výroku i s jeho negací ($BV \wedge B\neg V$), nemůže být dosažitelný žádný možný svět, protože v každém možném světě je splněna podmínka klasické logiky, že negace výroku je pravdivá právě tehdy, když výrok sám není pravdivý. V žádném možném světě tedy výroky V a $\neg V$ nemohou být současně pravdivé.

Řešení cvičení 30

Předpokládejme, že platí axiom T: $KA \Rightarrow A$. Verzi axiomu D dokážeme sporem. Nechť platí $K(A \wedge \neg A)$. Potom podle T také $A \wedge \neg A$, což je spor.

Nápověda ke cvičení 34

Zjistí, jak se liší množina možných světů dosažitelných z daného světa vzhledem k relacím \leq_R , \leq_S , \leq_V !

Řešení cvičení 34

Vzhledem k tomu, že vědci mají k dispozici více informací než sonda (například informace od dalších sond nebo informace z dalekohledů na zemi), považují méně možností za přípustné: méně

⁵⁸Řekneme, že relace dosažitelnosti je *reflexivní*.

stavů světa je v souladu s jejich poznatky. Naopak, všechny stavy světa, které vědci považují za možné (jsou dosažitelné v relaci \leq_V) jsou v souladu se všemi poznatky sondy (takže jsou dosažitelné i v relaci \leq_S). Jinak řečeno, jestliže pro dva světy P a Q platí $P \leq_V Q$, tak pro ně platí také $P \leq_S Q$ a pokud $P \leq_S Q$, tak $P \leq_R Q$. Pokud označíme množinu možných světů dosažitelných z nějakého pevně daného světa v relaci \leq_V (resp. \leq_S, \leq_R) symbolem M_V (resp. M_S, M_R), bude platit $M_V \subseteq M_S \subseteq M_R$.

