POZORUHODNÉ LOGICKÉ SYSTÉMY

Jaroslav Peregrin*
www.cuni.cz/~peregrin
[ORGANON F 8, 2000, 90-96, 210-217, 342-348, 460-466]

I
Hintikkova "logika podporující nezávislost"

V úvodním ročníku ORGANONu Pavel Cmorej ve svých *Kapitolách z logické syntaxe* předváděl, jak je přirozený jazyk možné nahlížet prismatem 'standardní' logiky. Historicky ovšem neexistuje jedna logika,ale různé logické systémy, které spolu částečně soupeří (tak jako třeba klasická a intuicionistická logika), částečně jeden druhý rozšiřují (jako třeba klasický výrokový a klasický predikátový počet) či se navzájem doplňují (jako například modální a temporální logika). To co je v logice obecně přijímáno za standard, je fakticky výsledkem interakce a soutěžení různých neustále vznikajících systémů.

V tomto čtyřdílném seriálu bych chtěl čtenářům ORGANONu přiblížit několik logických systémů, které byly vytvořeny či 'oprášeny' v nedávné době a jejichž vztah k logickým standardům je zatím ve stádiu diskuse. Jejich společným jmenovatelem je to, že jsou nějak zajímavé právě z hlediska analýzy přirozeného jazyka.

První ze systémů, kterému se budu věnovat, takzvaná *independence-friendly logic* (zkráceně IFL; "logika vstřícná nezávislosti")[1] navržená finským logikem a filosofem Jaakko Hintikkou, vzbudil v nedávné době poměrně velký rozruch. Jeho autor ho prezentoval s tvrzením, že *de facto* vyvrací slavnou Tarského větu stanovící, že žádný konzistentní jazyk není schopen vyjádřit svou vlastní pravdivost. Podívejme se nejprve na to, jak to s tímto vyvrácením je.

Paradox lháře zná každý: řeknu-li "Právě teď lžu", pak, jak se zdá, z předpokladu pravdivosti mého výroku vyplývá jeho nepravdivost a naopak. Obecně je tento paradox vyvoláván větou, která sama o sobě tvrdí, že je nepravdivá - to jest, můžeme říci, větou *Já nejsem pravdivá*. Jakmile mohu v nějakém jazyce takovou větu formulovat, mám paradox lháře a příslušný jazyk je tedy rozporný[2].

Co je potřeba k tomu, abych mohl v nějakém jazyce takovýto 'lhářovský výrok' formulovat? Zřejmě potřebuji (i) nějaký prostředek, který mi dovolí v rámci věty pojmenovat tuto větu samu ("já"), (ii) negaci ("ne"), a (iii) predikát pravdivosti ("být pravdivý"). Kdykoli tedy jazyk obsahuje tyto tři prostředky (a gramatickou možnost z nich sestavit inkriminovanou větu), je nutně nekonzistentní[3]. Tarského větu nyní můžeme vidět jako v podstatě triviální důsledek tohoto faktu: obsahuje-li jazyk negaci a mohou-li v něm výroky referovat k sobě sama, nemůže, nemá-li být sporný, obsahovat predikát pravdivosti.

Paradox je ovšem, teoreticky vzato, zřejmě možné blokovat i tak, že jazyku odepřeme některý jiný z potřebných prostředků. Poněkud problematické je to s (i), protože jak ukázal Gödel, jakmile je náš jazyk schopen vyjádřit elementární aritmetiku, je (i) triviálně splněn[4]. Mohli bychom ale jazyku samozřejmě odepřít negaci - a to je právě to, co dělá Hintikka. IFL tedy může obsahovat svůj vlastní predikát pravdivosti prostě proto, že neobsahuje negaci (přesněji řečeno obsahuje jenom 'negaci', která není aplikovatelná na každou větu).

Z tohoto pohledu tedy jeho 'vyvrácení Tarského' vypadá jako hodně laciný trik. Tak jednoduché to ale s Hintikkovou logikou přece jenom není - to, že IFL neobsahuje skutečnou negaci není jenom *ad hoc* tahem směřujícím k blokování paradoxu lháře; je výsledkem Hintikkových úvah o povaze logiky.

Hintikkova IFL vznikla na základě úvah o tzv. 'herní' (*game-theoretical*) sémantice pro standardní predikátový počet, kterým se tento autor věnoval po dlouhá desítiletí[5]. Takovou sémantiku můžeme vybudovat následujícím způsobem:
S každým (interpretovaným) výrokem V predikátového počtu spojíme hru H[V] pro dva hráče (které budeme nazývat 'Já' a 'Příroda') následujícím způsobem[6]:

I. H[$R(t_1,...,t_n)$] probíhá tak, že jsou-li $t_1,...,t_n$ v relaci R, vítězím *Já*, jinak vítězí *Příroda*

II. H[$\neg V$] probíhá tak, že si nejprve *Já* a *Příroda* vyměníme role, a pak se hraje H[V]

III. H[$V_1 \wedge V_2$] probíhá tak, že *Příroda* vybere jeden z výroků V_1, V_2 a hraje se jemu příslušní hra

IV. H[$V_1 \vee V_2$] probíhá tak, že *Já* vyberu jeden z výroků V_1, V_2 a hraje se jemu příslušná hra

V. H[$\exists x V[x]$] probíhá tak, že *Já* zvolím prvek a univerza a hraje se H[$V(a)$]

VI. H[$\forall x V[x]$] probíhá tak, že *Příroda* zvolí prvek a univerza a hraje se H[$V(a)$]

Tak například hra spojená s výrokem $\forall x \forall z\ \exists y\ \exists u\ R(x,z,y,u)$ probíhá následovně:

1. *Příroda* volí prvek x
2. *Příroda* volí prvek z.
3. *Já* volím prvek y.
4. *Já* volím prvek u.
5. Jsou-li x,z,y,u ve vztahu R, vítězím *Já*, jinak vítězí *Příroda*.

Nyní se dá snadno ukázat, že výrok V je pravdivý (v klasickém smyslu) právě tehdy, když mám *Já* vítěznou strategii ve hře H[V], a je nepravdivý právě tehdy, když má vítěznou strategii *Příroda*. Herní sémantika klasického predikátového počtu je tedy v tomto smyslu ekvivalentní sémantice standardní.

IFL nyní vznikne tak, že se v "herně" interpretovaném predikátového počtu připustí možnost 'utajování' předchozích tahů. To znamená, že se připustí i hry, při kterých nemá jeden z aktérů v momentě svého tahu dokonalou znalost o předchozích tazích svého protivníka. Hry tohoto nového typu budeme označovat pomocí nového druhu formulí využívajících symbolu "/" tak, že napíšeme-li T_2/T_1, bude to znamenat 'aktér tahu T_2 neví, jak jeho protihráč předtím provedl tah T_1'. Tak například hra spojená s výrokem $\forall x \forall z\ \exists y/\forall z\ \exists u/\forall x\ R(x,z,y,u)$ probíhá následovně:

1. *Příroda* volí prvek x.
2. *Příroda* volí prvek z.
3. *Já* volím prvek y, *aniž přitom vím, jaké* z *bylo zvoleno Přírodou.*
4. *Já* volím prvek u, *aniž přitom vím, jaké* x *bylo zvoleno Přírodou.*
5. Jsou-li x,z,y,u ve vztahu R, vítězím *Já*, jinak vítězí *Příroda*.

Vezměme například výrok (1) klasického predikátového počtu a výrok (2) IFL.

(1) $\forall x \forall z \exists y \exists u$ (**Obdivuje**$(x,y) \wedge$ **Obdivuje**$(z,u) \wedge (y \neq u)$)

(2) $\forall x \forall z\ \exists y/\forall z\ \exists u/\forall z$ (**Obdivuje**$(x,y) \wedge$ **Obdivuje**$(z,u) \wedge (y \neq u)$)

Představme si, že máme univerzum tvořené třemi individui, označme je A, B a C, a předpokládejme, že každé z těchto individuí obdivuje zbývající dvě. To znamená, že

A obdivuje B a C
B obdivuje A a C
C obdivuje A a B

Pak zřejmě platí, že výrok je (1) je pravdivý. Tento výrok je totiž očividně pravdivý právě tehdy, když si kterákoli dvě individua mohou zvolit reprezentanty svých 'obdivovanců' tak, aby se neshodla. (Tak například dvojice A a B to může udělat tak, že A zvolí B a B zvolí A.)

Jak to je s výrokem (2)? Ten je, řekli jsme, podle definice pravdivý právě tehdy, mám-li *Já* v příslušné hře vítěznou strategii; to znamená jsem-li na každý výběr x a z učiněný *Přírodou* schopen odpovědět takovým výběrem y a u, aby platilo (**Obdivuje**$(x,y) \wedge$ **Obdivuje**$(z,u) \wedge (y \neq u)$). Svoji volbu y ovšem mohu provést jenom na základě znalosti toho, jaké x zvolila *Příroda*, bez znalosti toho, jaké zvolila z; a podobně u musím zvolit jedině na základě znalosti z, nikoli x. To znamená, že nutnou podmínkou toho, abych měl vítěznou strategii, je existence takových přiřazení f a g prvků univerza prvkům univerza, aby když pro jakoukoli volbu x a z stanovím y jako $f(x)$ a u jako $g(z)$, bude platit (**Obdivuje**$(x,y) \wedge$ **Obdivuje**$(z,u) \wedge (y \neq u)$). Jinými slovy, (2) je pravdivý právě tehdy, když

(2') $\exists f \exists g \forall x \forall z$ (**Obdivuje**$(x,f(x)) \wedge$ **Obdivuje**$(z,g(z)) \wedge (f(x) \neq g(z))$)

To znamená, že (2) je pravdivý tehdy a jen tehdy, když existují dvě přiřazení obdivovaných obdivujícím, která mají disjunktní obory hodnot - a to v případě naší interpretace zjevně splněno není.

Výrok (2') budeme nazývat *skolemizací* výroku (2); a funkcím f a g ve (2') obsaženým budeme říkat Skolemovy funkce[7]. Abychom nahlédli, jaký podstatný rozdíl je mezi (1) a (2), uveďme příslušnou skolemizaci výroku (1):

(1') $\exists f\ \exists g\ \forall x \forall z$ (**Obdivuje**$(x,f(x,z)) \wedge$ **Obdivuje**$(z,g(x,z)) \wedge (f(x,z) \neq g(x,z))$)

Rozdíl je tedy v tom, že zatímco Skolemovy funkce obsažené v (2') jsou jednoargumentové, (1') vyžaduje *dvouargumentové* Skolemovy funkce.

Za jakých podmínek má ve hře odpovídající (2) vítěznou strategii *Příroda*? Zřejmě tehdy, když může zvolit taková dvě individua x a z, aby (**Obdivuje**$(x,y) \wedge$ **Obdivuje**$(z,u) \wedge (y \neq u)$) naplatila pro žádné y a u; a to zřejmě nastává právě tehdy když existují dvě individua, která si nemohou mezi svými obdivavanci vybrat tak, aby se neshodla. To však zřejmě v naší interpretaci opět neplatí - to je totiž splněno právě tehdy, když je výrok (1) nepravdivý (a (2) se tedy shoduje s (1) co do podmínek nepravdivosti; jakkoli se s ním neshoduje co do podmínek pravdivosti). Výrok (2) tedy není nepravdivý; a to znamená, že není, na rozdíl od (1), ani pravdivý, ani nepravdivý. (Příslušná hra je tedy vlastně, na rozdíl od hry příslušné kterémukoli výroku standardní logiky 1. řádu, 'férová' v tom smyslu, že ani jeden hráč není v předem beznadějné pozici.)

Podmínky pravdivosti (2) jsme, jak jsme viděli, schopni stanovit prostřednictvím jistého výroku predikátového počtu druhého řádu (totiž (2')). Dá se ukázat, že tohle platí obecně: pro každý výrok *V* IFL existuje výrok V^* klasického predikátového počtu 2. řádu (a to výrok typu Σ_1^1, to jest výrok tvořený řetězcem druhořádových existenčních kvantifikátorů následovaným prvořádovou formulí) tak, že *V* je pravdivý, právě když je V^* pravdivý. (Avšak neplatí, že *V* je nepravdivý, právě když je V^* nepravdivý!) Z tohoto faktu pak mimo jiné zjevně vyplývá, že výrok IFL má v IFL negaci, právě když je ekvivalentní výroku 1. řádu. (Důkaz: Buď *V* výrok IFL, který není výrokem klasické logiky 1. řádu. Předpokládejme, že má v IFL negaci *NV*. Buď *V'* formule predikátového počtu druhého řádu zachycující pravdivostní podmínky *V*. Pak negace *V* zachycuje pravdivostní podmínky *NV*. Avšak tato negace zřejmě není výrokem typu Σ_1^1 ani nemůže být žádnému takovému výroku ekvivalentní. Spor.)

Proč je IFL zajímavá? Bezesporu zajímavá, je z čistě formálního hlediska; má totiž některé pozoruhodné metalogické vlastnosti. Tou nejpozoruhodnější je samozřejmě její schopnost vyjádřit svůj vlastní predikát pravdivosti, o které jsme hovořili na počátku (jakkoli tato vlastnost, jak jsme naznačili, není zase *natolik* pozoruhodná, jak má tendenci prohlašovat Hintikka). Naznačme, jak může být v rámci IFL tento predikát definován (rigorózní definice je relativně složitá a je ji možné najít např. u Hintikky, 1997):

Jak ukázal Tarski, predikát pravdivosti pro standardní logiku můžeme definovat v jistém vhodném metajazyce (přičemž, a to je důsledkem Tarského věty zmiňované v úvodu tohoto pojednání, tento metajazyk musí být v podstatném smyslu 'bohatší' než jazyk objektový). Pro IFL je vnodné vzít za metajazyk predikátový počet 2. řádu: výše jsme naznačili, jak je možné prostřednictvím výroků této logiky vyjadřovat podmínky pravdivosti výroků jakékoli teorie v rámci IFL; a hledaná definice predikátu pravdivosti pro IFL je pak v podstatě konjunkcí obecných principů konstruování takových vyjádření[8]. Jak ovšem Hintikka ukazuje, bude tato definice výrokem typu Σ_1^1, a bude tak přeložitelná do IFL. Tak dostaneme vyjádření predikátu pravdivosti pro IFL v IFL (což je ovšem, jak jsme řekli na začátku, možné jenom díky tomu, že v IFL neexistuje kontradiktorická negace).

Přes tuto dramatickou odlišnost od klasického predikátového počtu si IFL zachovává některé příjemné vlastnosti predikátového počtu prvního řádu. Je totiž *kompaktní* (to jest jakákoli nekonečná teorie formulovaná v jazyce této logiky je konzistentní právě tehdy, když je konzistentní každá její konečná podmnožina); a navíc má i tzv. *Löwenheimovu-Skolemovu vlastnost* (teorie v jejím jazyce je splnitelná, je-li splnitelná ve struktuře se spočetným univerzem). To, čím IFL platí za zachování těchto 'příjemných' vlastností, je (sémantická) neúplnost: množina jejích platných formulí není rekurzivně vyčíslitelná[9].

Hintikka je ovšem přesvědčen, že IFL není zdaleka zajímavá jenom z tohoto formálního hlediska. Tvrdí totiž, že prostředky, které nám poskytuje a které chybí v klasické logice, jsou potřeba k adekvátnímu zachycení toho, co skutečně říkáme. Podle Hintikky má totiž taková věta jako

(3) Nějaký příbuzný každého vesničana a nějaký příbuzný každého měšťana se nenávidí,

alespoň jeden legitimní význam takový, že ho nelze zachytit jinak než prostřednictvím výroku IFL, totiž:

(3^*) $\forall x\ \forall z\ \exists y/\forall z\ \exists u/\forall x$ (**Vesničan**(x) $\wedge$ **Měšťan**(z) $\wedge$ **Příbuzný**(x,y) $\wedge$ **Příbuzný**(z,u) $\wedge$ **Nenávidí-se**(y,u))

Zvláště nezbytnou je podle Hintikky IFL pro adekvátní artikulaci toho, co říkáme v matematice: podle něj existuje celá řada zcela zásadních matematických pojmů, které nelze vyjádřit v rámci predikátové logiky 1. řádu. (Takové pojmy ovšem jistě lze vyjádřit v rámci logiky 2. řádu - Hintikka ovšem považuje IFL, z důvodů uvedených výše, za mnohem přijatelnější než plnou logiku 2. řádu.). Vezměme například pojem, který Frege (1884) vyjadřuje termínem "gleichzählig" (a o který se opírá při své explikaci pojmu čísla): jde o vztah, ve kterém jsou dvě vlastnosti právě tehdy, když je má stejný počet individuí, to jest když mají jejich extenze stejnou mohutnost. Dá se dokázat, že neexistuje žádná formule logiky 1. řádu, která by byla danou interpretací splňována právě tehdy, když se rovnají mohutnosti extenzí predikátů P a Q. Hintikka naproti tomu ukazuje, že formule IFL

(4) $\forall x\ \forall z\ \exists y/\forall z\ \exists u/\forall x\ ((\mathbf{P}(x) \rightarrow \mathbf{Q}(y)) \wedge (\mathbf{Q}(z) \rightarrow \mathbf{P}(u)) \wedge ((y{=}z) \leftrightarrow (u{=}x)))$

je splněna právě v tomto případě. Proč tomu tak je, je patrné z následující úvahy. Skolemizací (4) dostaneme formuli

(4') $\exists f\, \exists g\ \forall x \forall z\ ((P(x) \rightarrow Q(f(x)) \wedge (Q(z) \rightarrow P(g(z))) \wedge ((f(x){=}z) \leftrightarrow (g(z){=}x))),$

která, jak snadno nahlédneme, neříká nic jiného než to, že existuje vzájemně jednoznačné zobrazení extenze P na extenzi Q. (První konjunkt konstatuje, že f zobrazuje prvky extenze P na prvky extenze Q; druhý říká, že g zobrazuje prvky extenze Q na prvky extenze P, a třetí pak konstatuje, že f a g jsou vzájemně inverzní zobrazení.)

Tohle vede Hintikku k závěru, že je to právě IFL, co je tou nejvhodnější logikou jak pro analýzu přirozeného jazyka, tak pro budování základů matematiky: neustále dokonce hovoří o *revoluci v logice*[10]. My jsme ovšem upozornili na to, že jakkoli je IFL zajímavá, takováto prohlášení je třeba brát s jistou rezervou.

POZNÁMKY

[*] Práce na tomto textu byla podpořena grantem GA AV ČR číslo 401/99/0619.

[1] Viz Hintikka (1996a; 1997). S filosofickým pozadím Hintikkova přístupu měli čtenáři ORGANONu nožnost se seznámit prostřednictvím Kolářova překladu jedné z Hintikkových filosofičtěji orientovaných statí (viz Hintikka 1996b).

[2] Připomeňme, že právě tohle, i když v poněkud zakuklené podobě, ukázal v roce 1902 Bertrand Russell Fregovi o jazyce jeho logiky (viz Frege(1976)) . Russell totiž přišel na to, že v rámci Fregova systému je možné definovat pojem P tak, aby pro každý pojem p platilo $P(p) \leftrightarrow \neg p(p)$. Pojem P je tedy vlastně pojmem 'nespadání pod sebe sama' - $P(p)$ říká, že p nespadá pod sebe sama. V tomto smyslu tedy platí, že výrok $P(P)$ říká, že P nespadá pod sebe sama; tento výrok je ale současně zřejmě přímým konstatováním toho, že P pod sebe sama spadá - a vlastně tedy sám o sobě říká, podobně 'lhářovský výrok', že není pravdivý.

[3] Někdy se navíc uvádí, že je potřeba, aby v příslušném jazyce platila standardní logika - tento požadavek je ovšem podle mne obsažen již v požadavcích uvedených. Podle mého názoru nedává

smysl si představovat, že nějaký výraz by mohl znamenat totéž, co naše “ne” a přitom se neřídit zákony naší logiky.

[4] Gödel totiž ukázal, že pro každý (spočetný) jazyk nutně existuje jednoznačné přiřazení čísel všem výrazům; a je-li příslušný jazyk schopen vyjádřit aritmetiku, konkrétněji disponuje-li jmény pro přirozená čísla, můžeme právě tato jména brát jako pojmenování jeho výrazů.

[5] Viz například již Hintikka (1973).

[6] Několik poznámek k notaci, kterou budu zde i v následujících pokračováních užívat: Symboly s fixovanou interpretací píšu tučně, zatímco symboly, jejichž intepretace fixována není, píšu kurzívou. (To znamená, že kurzívou píšu jak proměnné, tak ale i ty extralogické konstanty, které jsou brány jako nespecifikované - takže píšu například “$\exists y \exists u$ **Obdivuje**(x,y)”, ale “pro nějaký predikát R platí $\exists x \exists y\, R(x,y)$”). Dále: $P(x)$ znamená predikát P aplikovaný na proměnnou x; zatímco $V[x]$ znamená výrok V obsahující proměnnou x; a následuje-li v jedné větě po symbolu $V[x]$ symbol $V[y]$, oznaèuje ten druhý, jak bývá zvykem variantu výroku V obsahující y tam, kde V obsahuje x.

[7] Upozorněme, že to není zcela standardní terminologie. Skolemizací (2) by se totiž obvykle nazývala nikoli (2’), což je výrok predikátového počtu 2. řádu, ale obdobná formule bez úvodních dvou existenčních kvantifikátorů.

[8] Definice pravdivosti pro IFL ovšem není dokonalou analogií Tarského definice pro standardní logiku. Vzhledem k tomu, že pro IFL obecně neplatí zákon vyloučení třetího, totiž zřejmě není možné pro každou teorii v rámci IFL definovat predikát pravdivosti **Pr** tak, aby z této definice pro každý výrok V a jeho Gödelovo číslo $\ulcorner V \urcorner$ vyplývala ekvivalence

$$\mathbf{Pr}(\ulcorner V \urcorner) \leftrightarrow V^*,$$

kde V^* je ‘stoprocentním’ překladem V do metajazyka. (Platnost této ekvivalence totiž zřejmě implikuje, že pro V^*, a potažmo pro V, platí zákon vyloučení třetího.) Co možné je, je definovat **Pr** tak, aby výše uvedená ekvivalence platila vždy v případě, že V^* je oním vyjádřením podmínek pravdivosti V v logice druhého řádu, o jakých jsme hovořili výše. V každém případě pak bude $\mathbf{Pr}(\ulcorner V \urcorner)$ platit tehdy a jen tehdy, když je $\ulcorner V \urcorner$ Gödelovým číslem pravdivého výroku - takže dává i přes tuto odchylku stále dobrý smysl mluvit o definici *pravdivosti*. Hledaná definice je pak analogická klasické Tarského definici.

[9] Čtenář, který zná tzv. Lindströmovu větu (viz Lindström, 1969), se může podivovat, zda je tohle skutečně možné - Lindströmova věta totiž říká, že klasický predikátový počet nelze netriviálně rozšířit, aniž bychom tím přišli buď o kompaktnost nebo o Löwenheimovu-Skolemovu vlastnost. K tomu je třeba si uvědomit, že IFL není rozšířením klasické logiky v Lindströmově smyslu, protože neobsahuje klasickou negaci. Jestliže k ní negaci přidáme, jak o Löwenheimovu-Skolemovu vlastnost, tak o kompaktnost přijdeme.

[10] Hintikkovu argumentci v Čechách převzal Jiří Fiala (1997). Podrobnější kritiku tohoto stanoviska lze nalézt v mé polemické reakci na Fialův článek (Peregrin, 1998). Viz též polemické reakce Hájka a Sochora (1998) a Hájka (1998).

LITERATURA:

[1] FIALA, J. (1997): Je elementární logika totéž co logika prvního řádu? In: **Pokroky matematiky fyziky a astronomie** 42, 127-133.

[2] FREGE, G. (1884): **Grundlagen der Arithmetik**. Koebner, Breslau.

[3] FREGE, G. (1976): **Wissenschaftlicher Briefwechsel** (ed. G. Gabriel et al.). Meiner, Hamburg.

[4] HÁJEK, P. a SOCHOR, A. (1998): Klasická logika v kontextu svých zobecnění. **Pokroky matematiky fyziky a astronomie** 43, 39-45.
[5] HÁJEK, P. (1998): Ještě o elementární logice. **Pokroky matematiky fyziky a astronomie** 43, 324-325.
[6] HINTIKKA, J. (1973): **Logic, Language-Games and Information**. Clarendon Press, Oxford.
[7] HINTIKKA, J. (1996a): Contemporary Philosophy and the Problem of Truth. **Acta Philosophica Fennica** 61 (**Methods of Philosophy and the History of Philosophy**, ed. S. Knuttila a I. Niiniluotto). Èeský pøeklad **Organon F** 4 (1997), 137-154.
[8] HINTIKKA, J. (1996b): **The Principles of Mathematics Revisited**. Cambridge University Press, Cambridge.
[9] HINTIKKA, J. (1997): **Lingua Universalis vs. Calculus Ratiocinatur** (**Selected papers**, vol. 2). Kluwer, Dordrecht. (Zvláště článek Defining Truth, the Whole Truth and Nothing But Truth původně publikovaný v r. 1991)
[10] LINDSTRÖM, P. (1969): On Extensions of Elementary Logic. **Theoria** 35, 1-11.
[11] PEREGRIN, J (1998): Co je to elementární logika? **Pokroky matematiky fyziky a astronomie** 43, 45-47.

II.
Hilbertův epsilon-kalkul a současné pokusy o jeho využití pro analýzu jazyka

Epsilon kalkulus je logický systém, jehož historie je dlouhá; navrhl ho již mezi dvěma světovými válkami David Hilbert. Od té doby upadl bezmála do zapomnění, až byl v nedávné době oprášen některými logiky, kteří se zabývají analýzou přirozeného jazyka. Podívejme se nejprve na původní Hilbertovu verzi.

Představme si, že k jazyku standardního predikátového počtu prvního řádu přidáme následující gramatické pravidlo vytvářející termy z formulí[1]:

(1) je-li F formule a x proměnná, je εxF term

Přijetí tohoto pravidla bude zřejmě znamenat, že kromě standardních termů, jakými mohou být třeba **Karel,** x či **otec**(**Karel**) budou moci být správně utvořenými termy například i εx**Filosof**(x) či εx(**Karel** = **otec**(x)); a tudíž správně utvořenými výroky budou i například **Filosof**(εx(**Karel** = **otec**(x))) či **Karel** = εx**Filosof**(x). Představme si dále, že k standardním axiomům predikátového počtu přidáme schéma

(2) $F[T] \rightarrow F(\varepsilon xF[x])$

To v podstatě říká, že má-li nějaké individuum vlastnost vyjadřované formulí $F[x]$, má tuto vlastnost určitě i to individuum, které je označováno termem $\varepsilon xF[x]$. (V logice se někdy pro individua tohoto druhu užívá termín *scapegoat*, 'obětní beránek' - jeho smysl se stane jasným, když si představíme, že $F[x]$ vyjadřuje nějakou nezáviděníhodnou vlastnost, třeba *být zbit*. $\varepsilon xF[x]$ je pak, můžeme říci, ten nebožák, kterého zaručeně zbijí, kdykoli zbijí vůbec někoho[2].) Vzhledem k tomuto si jako 'zamýšlenou interpretaci' termu $\varepsilon xF[x]$ pracovně představit něco jako 'nějaký F': takže třeba εx**Filosof**(x) čteme jako 'nějaký filosof' a **Karel** = **otec**(εx**Filosof**(x)) jako 'Karel je otcem nějakého filosofa' (dále ovšem uvidíme, že takové čtení může být i zavádějící; usnadní však prvotní orientaci).

Snadno se ukáže, že takto definovaný epsilon kalkulus (kterému se z důvodů, které vyjdou najevo za chvíli, říká *intenzionální* epsilon kalkulus) je konzervativním rozšířením standardního predikátového počtu (to znamená, že přidáním (1) a (2) se nijak nezmění logické vlastnosti toho, co už v predikátovém počtu bylo před ním). Navíc se snadno ukáže, že za předpokladu, že univerzum neobsahuje nepojmenované objekty, platí následující dvě ekvivalence:

(3) $\exists xF[x] \leftrightarrow F(\varepsilon xF[x])$
(4) $\forall xF[x] \leftrightarrow F(\varepsilon x\neg F[x])$

Že platí (3) je zřejmé: přímá implikace vyplývá přímo z (2); nepřímá je instancí pravidla existenciální instanciace. (4) se z (3) dostane dosazením $\neg F$ za F a znegováním obou stran ekvivalence. Za uvedeného předpokladu se tedy nabízí možnost přijmout (3) a (4) nikoli jako teorémy, ale jako *definice* standardních kvantifikátorů.

Protože operátor ε vytváří termy z predikátů, dalo by se očekávat, že bude vyjadřovat funkci, přiřazující individua třídám individuí - tak by tomu bylo ovšem jenom tehdy, kdybychom měli zaručeno, že $\varepsilon x F[x]$ a $\varepsilon x G[x]$ označují totéž individuum, kdykoli F a G označují tutéž třídu individuí; tedy že platí

(5) $\forall z(F[z] \leftrightarrow G[z]) \rightarrow (\varepsilon x F[x] = \varepsilon x G[x])$

(5) je vlastně jakýsi požadavek 'extenzionality': zaručuje, že individuum, které je označováno $\varepsilon x F[x]$, je určeno jedině extenzí $F[x]$. A právě to, že (5) není v rámci výše definovaného epsilon-kalkulu obecně splněno, je důvodem, proč se tento kalkulus nazývá *intenzionální*. Jestliže k němu (5) přidáme jako další axiom, dostaneme *extenzionální epsilon kalkulus*.

Jakou sémantikou bychom mohli epsilon kalkulus opatřit? Intuitivně se to zdá být zřejmé: $\varepsilon x F(x)$ označuje individuum (a sice takové individuum, které patří do množiny označované predikátem F), takže ε by měl, jak už jsme řekli, reprezentovat funkci, která přiřazuje individua množinám individuí. Avšak kterou funkci? Funkcí zobrazující každou neprázdnou podmnožinu univerza na nějaký její prvek je jistě mnoho. A ukazuje se, že jediným řešením je prostě relativizace sémantiky epsilon kalkulu k volbě této funkce. To znamená, že zatímco u standardního predikátového počtu jsou výrazům přiřazovány denotáty v závislosti na interpretaci extralogických konstant a na valuaci proměnných, v případě epsilon kalkulu musíme přidat další parametr, výběrovou funkci.

Je-li tedy I interpretace extralogických konstant, V valuace proměnných a Φ výběrová funkce přiřazující každé podmnožině univerza prvek univerza (je-li neprázdná, pak *její* prvek) a označíme-li symbolem $\mathrm{V}[x|\mathrm{i}]$ tu valuaci proměnných, která je jako V s tím jediným možným rozdílem, že proměnné x přiřazuje individuum i, můžeme definovat denotování termů a splňování formulí simultánní indukcí[3]:

$\|T\|_{\mathrm{I},\Phi,\mathrm{V}} = \mathrm{I}(T)$ pro každý konstantní term T
$\|x\|_{\mathrm{I},\Phi,\mathrm{V}} = \mathrm{V}(x)$ pro každou proměnnou x
$\|f(t_1...t_n)\|_{\mathrm{I},\Phi,\mathrm{V}} = \mathrm{I}(f)(\|t_1\|_{\mathrm{I},\Phi,\mathrm{V}},..,\|t_n\|_{\mathrm{I},\Phi,\mathrm{V}})$, pro každý funktor f a termy $t_1...t_n$
$\|\varepsilon x F\|_{\mathrm{I},\Phi,\mathrm{V}} = \Phi(\{\mathrm{i} : \mathrm{I},\Phi,\mathrm{V}[x|\mathrm{i}] \models F\})$,
$\mathrm{I},\Phi,\mathrm{V} \models P(t_1...t_n)$, jestliže $<\|t_1\|_{\mathrm{I},\Phi,\mathrm{V}},..,\|t_n\|_{\mathrm{I},\Phi,\mathrm{V}}> \in \mathrm{I}(P)$
$\mathrm{I},\Phi,\mathrm{V} \models F_1 \wedge F_2$, jestliže $\mathrm{I},\Phi,\mathrm{V} \models F_1$ a $\mathrm{I},\Phi,\mathrm{V} \models F_2$
atd.

$\mathrm{I,s} \models F$ pak definujeme jako platné právě tehdy, když $\mathrm{I},\Phi,\mathrm{s} \models F$ pro každou výběrovou funkci Φ[4].

Vezměme například term εx**Filosof**(x). Podle definice $\|\varepsilon x\mathbf{Filosof}(x)\|_{\mathrm{I},\Phi,\mathrm{V}} = \Phi(\{\mathrm{i} : \mathrm{I},\Phi,\mathrm{V}[x|\mathrm{i}] \models \mathbf{Filosof}(x)\})$; to znamená, že denotát tohoto termu je hodnotou výběrové funkce Φ aplikované na množinu těch individuí, která splňují formuli **Filosof**(x), to jest na extenzi predikátu **Filosof**, neboli množinu všech filosofů. Podobně bude denotátem termu εx(**Karel** = **otec**(x)) 'vybraný Karlův syn' – to jest to individuum, které výběrová funkce Φ přiřazuje množině všech Karlových synů. (Přitom si uvědomme, že v případě, že Karel žádné syny nemá, bude denotátem termu εx(**Karel** = **otec**(x)) individuum, které Karlovým synem *není*. Bylo by samozřejmě přirozenější, aby v tomto případě neměl tento term denotát žádný – vzhledem k tomu, že epsilon kalkulus je budován jako nadstavba standardní predikátové logiky, která nedonotující termy nepřipouští, však tato možnost není k dispozici.)

Myšlenka, se kterou původně Hilbert tento kalkul navrhl, byla redukce axiomů kvantifikace na axiom výběru. To se postupně přestalo jevit jako zajímavý cíl, a espilon kalkulus tak upadl téměř v zapomnění. V poslední době však došlo k jeho pozoruhodné resuscitaci, a to v důsledku toho, že byl shledán zajímavým lidmi, kteří se zabývají analýzou přirozeného jazyka. Důvodem bylo především to, že se Hilbertův epsilon-operátor začal jevit jako perspektivní z hlediska analýzy anglického neurčitého členu.

Abychom tohle vysvětlili, vraťme se na okamžik ještě před Hilberta, k takzvanému iota(-inverzum[5])-operátoru, navrženému Bertrandem Russellem. Russell sám tento operátor zavedl jako kontextuálně definovanou zkratku[6]:

(6) $F[\iota x G[x]] \equiv_{\mathrm{Def}} \exists x(F[x] \wedge G[x] \wedge \forall y(G[y] \rightarrow (x = y)))$,

takže například výrok, analyzující Russellovu oblíbenou větu

(7) Král Francie je holohlavý,

totiž

(7′) $\exists x$ (**Král-Francie**(x) $\wedge$ **Holohlavý**(x) $\wedge$ $\forall y$(**Král-Francie**(y) $\rightarrow$ ($x = y$))),

bude zapsán krátce jako

(7″) **Holohlavý**(ιx**Král-Francie**(x)).

Připustíme-li však možnost termů, které neoznačují nic (to jest přejdeme-li k logice, která připouští modely s parciálními funkcemi), můžeme iota-termy přirozeným způsobem pojmout jako termy plnohodnotné. Analogicky jako v případě epsilon-termů totiž můžeme přidat syntaktické pravidlo

(8) je-li F formule a x proměnná, je $\iota x F$ term,

které je v tomto případě možné ošetřit sémanticky zcela přímočaře:

$\|\iota x\mathrm{F}\|_{\mathrm{I,V}} = \mathrm{i}$, je-li i jediným prvkem množiny $\{\mathrm{i} : \mathrm{I,V}[x|\mathrm{i}] \models F\}$
a není definováno ve všech ostatních případech.

To znamená, že na ι se můžeme dívat jako na reprezentaci funkce, která přiřazuje jednoprvkové množině její jediný prvek a kterékoli jiné množině nic. ι se tak zdá být vhodnou logickou explikací anglického určitého členu: *president of the USA* [*prezident USA*] je predikátem, jehož extenzí je jednoprvková množina tvořená v současné době Billem Clintonem, a *the president of the USA* [*ten jediný prezident USA*] je jménem tohoto individua; zatímco *king of France* [*král Francie*] resp. *Olympic winner* [*olympijský vítěz*] mají za extenzi množinu prázdnou resp. více než jednoprvkovou, a *the king of France* [*ten jediný král Francie*] resp. *the Olympic winner* [*ten jediný olympijský vítěz*] jsou proto ‘jména’, která (momentálně) neoznačují nic[7].

Nyní se může zdát, že epsilon-operátor by mohl posloužit jako logický protipól *neurčitého* členu analogicky tomu, jako jsme právě iota-operátor užili jako protipól členu určitého. Ostatně jsme řekli, že je-li $\iota x P(x)$ intuitivně chápáno jako ‘to jediné P’, zamýšlenou interpretací $\varepsilon x P(x)$ má být ‘nějaké P’. Je ovšem třeba si uvědomit, že tato analogie, tak jak je, pokulhává: ε totiž nemá vlastní sémantiku v tom smyslu, v jakém ji má ι. ι označuje *jednu určitou* (parciální) funkci z množin individuí do individuí; avšak totéž, jak jsme viděli, o ε říci nelze[8].

Klaus von Heusinger (1997) se pokusil tento nedostatek odstranit tím, že ‘podspecifikovanost’ sémantiky ε prohlásil za záležitost pragmatiky. Sémanticky je podle něj dáno to, že ε označuje *nějakou* výběrovou funkci přiřazující neprázdným množinám jejich prvky, a kterou konkrétně, je dáno až kontextem příslušné promluvy. Podle tohoto názoru je například fráze *a philosopher* [*nějaký filosof*] mnohoznačná, a zjednoznační se až v určitém kontextu: řeknu-li *I saw a philosopher yesterday* [*Včera jsem viděl nějakého filosofa*], omezím tím množinu ‘přijatelných’ výběrových funkcí na ty, které přiřazují množině filosofů toho, kterého jsem já včera viděl.

Výraz $\varepsilon x\mathrm{F}[x]$ tedy, můžeme říci, označuje konkrétní individuum jedině tehdy, je-li specifikována jedna z mnoha možných výběrových funkcí (či je-li alespoň specifikována nějaká vhodná podmnožina takových funkcí). V rámci formální definice sémantiky epsilon-kalkulu je tato specifikace dána funkcí Φ (která se tak ovšem stává dalším parametrem sémantické interpretace); v rámci přirozeného jazyka je pak takový výběr, podle von Heusingerova návrhu, ustanovován kontextem.

Co by tedy pak bylo *významem* neurčitého členu? Je jím jedna konkrétní výběrová funkce? Jistě ne: neurčitý člen se chová tak, že je-li kontextem stanovena výběrová funkce, pak dané množině přiřadí jejího reprezentanta. Ač Heusinger sám to takto neformuluje, můžeme se na takto uchopený neurčitý člen dívat jako na denotující funkci, která dané výběrové funkci přiřadí funkci, která množině objektů přiřadí jejího reprezentanta. Je-li tedy m libovolná množina univerza a Φ libovolná výběrová funkce, můžeme psát

$$(\|\varepsilon\|(\Phi))(m) = \Phi(m).$$

Z toho ovšem zřejmě vyplývá

$$\|\varepsilon\|(\Phi) = \lambda m.\Phi(m) = \Phi,$$

a tedy

$$\|\varepsilon\| = \lambda\Phi.\Phi$$

Dospíváme tedy k závěru, že ε, a potažmo neurčitý člen, označuje z tohoto úhlu pohledu identické zobrazení množiny výběrových funkcí na sebe sama. (Operátor ε je totiž nahlížen jako denotující, podobně jako ι, výběrovou funkci, ale na rozdíl od ι ne výběrovou funkci pevně danou, nýbrž tu výběrovou funkci, která je právě ‘dodávána’ kontextem. Neurčitý člen tedy z tohoto pohledu nedělá nic jiného, než že ‘vyzvedává’ z kontextu aktuální výběrovou funkci, kterou pak aplikuje na extenzi predikátu, na nějž je aplikován. Tak například v rámci termu εx**Filosof**(x) funguje ε z tohoto pohledu tak, že nejprve ‘převezme’ z kontextu aktuální výběrovou funkci a tu pak aplikuje na extenzi predikátu **Filosof**.)

Domnívám se, že takovýto pohled má dva nedostatky. Zaprvé, skutečně jasný smysl může dát až v kontextu nějakého logického systému, v jehož rámci lze pracovat s ‘kontexty’ – to jest až v rámci nějaké *dynamické* verze sémantiky, o jakých budeme hovořit v příštím pokračováním tohoto seriálu (viz též Peregrin, 1996). V rámci této sémantiky jsou totiž významy výroků, a potažmo jejich částí chápány jako zobrazení množiny ‘kontextů’ (či ‘informačních stavů’) na tutéž množinu - a jak uvidíme, je výběrové funkce možné vidět jako způsoby specifikace právě určitých relevantních rysů kontextů. Zadruhé, a to je podstatnější, když pak podrobněji analyzujeme to, jak se neurčitý člen fakticky chová, dojdeme k závěru, že ve skutečnosti *mění* kontext: To vyplývá z toho, že například věta *I saw the philosopher too* [*Taky jsem toho filosofa viděl*], která nedává smysl v ‘prázdném kontextu’, smysl dává v kontextu ‘vyrobeném’ třeba větou *Yesterday I saw a philosopher in a funny green jacket* [*Včera jsem viděl filosofa v legračním zeleném kabátku*]. (Tato ‘dynamická’ analýza pak vede také k adekvátnější analýze určitého členu, než je ta russellovská - ale o tom skutečně až příště.)

Vraťme se nyní ke klasické, neparciální verzi epsilon-kalkulu, kterou jsme uvedli na počátku tohoto pojednání. V této podobě má epsilon kalkulus jednu bizarní vlastnost, která dělá z ε něco zásadně jiného, než je neurčitý člen. Protože v této verzi epsilon kalkulu nejsou připuštěny nedenotující termy, musí $\varepsilon x F[x]$ něco denotovat i tehdy, když neexistuje nic, co by splňovalo formu $F[x]$. Výrok tvaru $G[\varepsilon x F[x]]$ tedy *vždy* připisuje vlastnost nějakému individuu: tak výrok **Holohlavý**(εx**Král-Francie**(x)) připisuje i za současného stavu věcí holohlavost nějakému individuu (které ovšem samozřejmě není králem Francie).

To se zdá být přinejmenším podivné; Hartley Slater (1984) se však pokusil tuto podivnost interpretovat jako zásadní *přednost* epsilon-kalkulu. Tento rys tohoto kalkulu totiž podle něj způsobuje, že rozšířením predikátového počtu na epsilon kalkulus se dostáváme do oblasti *intenzionální* logiky (teď se ovšem myslí *intenzionalita* v tom běžném slova smyslu, v jakém je intenzionální logikou třeba TIL, ne onen formální, dříve zmiňovaný smysl).

Abychom vysvětlili, jak to Slater myslí, uvažme Fregovu oblíbenou větu

(9) Jitřenka je Večernice.

Takový výrok je podle Slatera v rámci epsilon-kalkulu analyzovatelný dvěma způsoby, totiž

(9′) $\exists x\,(\mathbf{Jitřenka}(x) \wedge \forall y\,(\mathbf{Jitřenka}(y) \rightarrow (x = y)) \wedge$
$\mathbf{Večernice}(x) \wedge \forall y\,(\mathbf{Večernice}(y) \rightarrow (x = y)))$,

což bychom mohli pomocí Russellovy definice (6) zapsat jako

$$\iota x\mathbf{Jitřenka}(x) = \iota x\mathbf{Večernice}(x),$$

a vedle toho jako

(9″) $\varepsilon x(\mathbf{Jitřenka}(x) \wedge \forall y\,(\mathbf{Jitřenka}(y) \rightarrow (x = y)) =$
$\varepsilon x(\mathbf{Večernice}(x) \wedge \forall y\,(\mathbf{Večernice}(y) \rightarrow (x = y)))$

(9′) a (9″) přitom nejsou ekvivalentní: zatímco první z nich je pravdivý jenom kontingentně, ten druhý je identitou a jako takový nemůže být v rámci klasické logiky kontingentní, konkrétně vzhledem k tomu, že je aktuálně pravdivý, musí být pravdivý nutně. Podle Slatera tak epsilon-kalkulus dovede vyjádřit jak extenzionální, tak intenzionální úroveň sémantické analýzy identity: (9′) podle něj konstatuje (kontingentní) identitu extenzí, zatímco (9″) vyjadřuje nutnou ekvivalenci intenzí.

Jiným dokladem intenzionality epsilon-kalkulu je podle Slatera fakt, že v něm můžeme hovořit o neexistujících objektech. Vezměme větu

(10) Fajst se domnívá, že potkal ducha,

převedeno do jazyka extenzionální logiky[9]

(10′) **Domnívá-se**(**Fajst**,$\exists x$(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x))).

Protože v epsilon-kalkulu má každý term zaručeně denotát, mají denotát i následující termy:

(11) εx**Potkal**(**Fajst**,x)
(12) εx(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x)),
(13) εx**Domnívá-se**(**Fajst**,(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x))),

To podle Slatera znamená, že nám epsilon kalkulus dovoluje mluvit i o neexistujících, 'intenzionálních' objektech – třeba o 'duchovi, kterého potkal Fajst' či o 'tom, o čem se pan Fajst domnívá, že je to duch, kterého potkal'.

Všimněme si některých dalších pozoruhodných rysů takových logických analýz. Výrok

(14) $\neg$**Duch**(εx(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x)))

ba dokonce ani výrok

(15) $\neg$**Duch**(εx(**Duch**(x)))

není kontradiktorická. (Všimněme si, že kdybychom se drželi našeho čtení $\varepsilon xF[x]$ jako 'nějaký F', museli bychom (15) číst jako 'nějaký duch není duch' a vzhledem k tomu, že (15) není kontradikce, bychom museli připustit, že 'nějaký duch nemusí být duch'.) Jestliže totiž duchové neexistují, je (15) podle (4) pravdivá. Kontradiktorická tedy není ani formule

(16) εx**Potkal**(**Fajst**,x) = **Kus-hořícího-karbidu**

Předpokládejme, že je (16) pravda. Protože z (10′) můžeme zřejmě odvodit

(17) **Domnívá-se**(**Fajst**,**Duch**(εx(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x))))

a z toho pak dále

(18) **Domnívá-se**(**Fajst**,**Duch**(εx(**Potkal**(**Fajst**,x)))).

dostáváme jako důsledek (10′) a (16)

(19) **Domnívá-se**(**Fajst**,**Duch**(**Kus-hořícího-karbidu**)).

Toto zjevně není nic jiného, než výsledek neopodstatněného nahrazování koextenzionálních termínů v neextenzionálním kontextu. Podle Slatera pak ovšem v rámci epsilon-kalkulu tou 'správnou' identitou, kterou použít lze, je nikoli (16), ale (20):

(20) εx**Domnívá-se**(**Fajst**,(**Duch**(x) $\wedge$ **Potkal**(**Fajst**,x))) = **Kus-hořícího-karbidu**.

S *její* pomocí totiž (19) neodvodíme.

Mně osobně se zdá, že klíčem k pochopení skutečného dosahu Slaterovy ekvilibristiky je pochopení faktu, že epsilon-termy jsou sémanticky zcela absurdní. Představme si, že bych definoval term **Bimbo**, který by v sudé dny denotoval slona Bimba, zatímco v liché dny třeba nějakou žábu; a pak bych z faktu, že výrok **zelený**(**Bimbo**) denně mění pravdivostní hodnotu, vyvozoval, že je Bimbo vlastně chameleon. Podobně je to s epsion termy: $\varepsilon xF[x]$ 'normálně' denotuje 'nějakého F'; avšak za specifických okolností (když žádný F neexistuje) denotuje i něco jiného, a to vyvolává iluzi, jako by F někdy nemusel být F. Faktem ovšem zůstává, že v podstatě není možné najít renomovaný časopis z oboru filosofické logiky, který by výše uvedené Slaterovy výsledky nebyl otiskl.

POZNÁMKY:

[1] I v tomto pokračování používám stejné notační konvence jako v předchozí části: to znamená, že symboly s fixovanou interpretací píši tučně, zatímco symboly, jejichž intepretace fixována není, uvádím kurzívou; dále $P(x)$ znamená predikát P aplikovaný na proměnnou x, zatímco $F[x]$ znamená formuli F obsahující proměnnou x; a následuje-li v jedné větě po symbolu $F[x]$ symbol $F[y]$, označuje ten druhý variantu výroku F obsahující y tam, kde F obsahuje x.

[2] Samozřejmě že tento vtip je založen na zavádějící představě, že příslušný term má nějakou pevnou interpretaci napříč modely.

[3] Zde je notační konvence taková, že $\|T\|_{\mathrm{I},\Phi,\mathrm{V}}$ označuje denotát termu T vzhledem k interpretaci I, výběrové funkci Φ a valuaci V; zatímco zápis I,Φ,V $\models F$ říká, že formule F je splňována (či že je pravdivá vzhledem k) I, Φ a V.

[4] Podrobný rozbor logických vlastností epsilon kalkulu podává Meyer Viol (1995).

[5] Russell původně používal řecké písmeno iota obrácené vzhůru nohama. Vzhledem k tomu, že taková věc je pro sazeče noční můrou, setkáváme se dnes již běžně s používáním obyčejného, nepřevráceného iota.

[6] Viz Russell a Whitehead (1910, §14).

[7] Viz též Cmorej (1995)

[8] Viz též Peregrin (2000).

[9] Tady jsme už ovšem za hranicí logiky prvního řádu – uvažujeme totiž predikát, který má za argument výrok.

LITERATURA:

[1] CMOREJ, P. (1995): Z logickej syntaxe a sémantiky VII. **ORGANON F** 2, 306-318.
[2] HILBERT, D. a BERNAYS, P. (1934; 1939): **Grundlagen der Mathematik**. Springer, Berlin.
[3] MEYER-VIOL, W.P.M. (1995): **Instantial Logic** (dissertation). ILLC, Amsterdam.
[4] PEREGRIN, J. (1996): Dynamická sémantika. **ORGANON F** 4, 333-348.
[5] PEREGRIN, J. (2000): Reference and Inference: the Case of Anaphora. In: **Reference and Anaphoric Relations** (ed. K. von Heusinger a U. Egli). Kluwer, Dordrecht, 269-286.
[6] RUSSELL, B. a WHITEHEAD, A.N. (1910): **Principia Mathematica I**. Cambridge University Press, Cambridge.
[7] SLATER, B.H. (1994): **Intensional Logic**. Gower, Aldershot.
[8] VON HEUSINGER, K. (1997): **Salienz und Referenz**. Akademie, Berlin.

III.
Dynamická logika

Uvažme výroky

(1) Možná prší
(2) Neprší

Teorie tvořená těmito dvěma výroky je, měřeno běžnými logickými standardy, zjevně konzistentní. Znázorněno prostředky modální či intenzionální logiky výrok (1) říká, že existuje nějaký možný svět (dosažitelný z toho našeho, pokud pracujeme s kripkovskou relací dosažitelnosti), ve kterém prší; zatímco (2) říká, že tímto světem není přímo ten náš. To zřejmě není obecně v rozporu.

Porovnejme ale to, když někdo konstatuje

(3) Možná prší. Neprší.

s tím, když prohlásí.

(4) Neprší. Možná prší.

Intuitivně je mezi těmito dvěma vyjádřeními rozdíl: máme pocit, že v tom druhém případě si mluvčí jaksi odporuje; že zatímco tvrzením *Neprší* se možnost následného tvrzení *Možná prší* uzavírá, prohlásíme-li nejprve *Možná prší*, neuzavřeme si tím možnost následně prohlásit *Neprší*. Jak se to srovnává se standardní logickou analýzou, z jejíhož hlediska by měly být tyto výroky slučitelné bez ohledu na pořadí, ve kterém za sebou následují?

Intuice, o kterou jde, zřejmě vyplývá z toho, že onomu „možná" v (1) běžně rozumíme poněkud jinak, než jak to explikuje standardní intenzionální logika. Výroku (1) totiž nerozumíme jako konstatování, že pršení není vyloučeno logicky, ale spíše jako konstatování, že není vyloučeno tím, co dosud víme – to jest že nějaký možný svět, ve kterém prší, nejenom prostě existuje či je dosažitelný z toho našeho, ale že patří mezi ty světy, které jsou slučitelné se současným stavem našich vědomostí. Toto „možná" má tedy *epistemický* smysl, který není prostředky intenzionální logiky dost dobře uchopitelný: vyjadřuje totiž v podstatě to, že náš současný stav vědomostí nevylučuje výrok, na který je aplikováno.

Jak tohle vysvětluje rozdíl mezi (3) a (4)? Konstatujeme-li (1) v rámci (3), to jest *po* konstatování (2), konstatujeme ho zřejmě v kontextu jiného stavu vědomostí, než když ho konstatujeme v rámci (4), tedy *před* konstatováním (2). (2) totiž zřejmě nese netriviální informaci, a v důsledku toho jeho konstatování může způsobit rozšíření našich vědomostí.

Abychom tohle mohli analyzovat formálně, musí být naše sémantická teorie nějak schopna zachytit, jak se stav našich vědomostí vyvíjí v průběhu diskurzu. (Všimněme si, že ač se má obvykle za to, že sémantiku je třeba nesměšovat s epistemologií, tady jsme k epistemologickým úvahám vedeni přímo snahou o adekvátní explikaci sémantiky některých výrazů našeho jazyka.) K jednoduché verzi takového modelu ale není těžké dospět od standardního intenzionálního modelu jazyka. ‘Stav vědomostí‘ lze totiž zřejmě zachytit jako určitou množinu propozic; a protože množina propozic je z daného hlediska ztotožnitelná s konjunkcí svých prvků, můžeme stav vědomostí explikovat jako propozici, tedy v rámci intenzionální logiky jako množinu možných světů, konkrétně jako množinu všech těch možných světů, které jsou slučitelné s tím, co dosud víme.

Podle tohoto modelu můžeme konstatování jednoduché propozice, jako je (2), vidět jako záležitost vylučování možných světů. Konstatuji-li (2), rozšiřuji tím daný informační stav tak, že zužuji jej reprezentující množinu možných světů na ty, ve kterých neprší. Takové propozici, jako je (2), pak mohu přiřadit něco, čemu lze říkat *potenciál změny informačního stavu* (*PZIS*)[1] a co lze explikovat jako funkci, přiřazující informačnímu stavu (tedy množině možných světů) informační stav (množinu možných světů). Mezi PZIS výroku a jeho intenzí bude ovšem jednoduchý vztah: PZIS výroku V je funkce, která dané množině M možných světů přiřadí průnik této množiny s intenzí V, to jest označíme-li intenzi V jako $\|V\|^{\mathbf{I}}$ a PZIS V jako $\|V\|^{\mathbf{P}}$, bude platit

$$\|V\|^{\mathbf{P}}(\mathrm{M}) = \mathrm{M} \cap \|V\|^{\mathbf{I}},$$

to jest

$$\|V\|^{\mathbf{P}} = \lambda m\, (m \cap \|V\|^{\mathbf{I}}).$$

Užití výroku v některém kontextu ovšem může vyústit v jakousi ‘informační frustraci’ - v případě některých V bude $\|V\|^{\mathbf{P}}(\mathrm{M})$ pro některé neprázdné množiny M množinou prázdnou. (Taková množina bude zřejmě existovat pro každý výrok, který neplatí nutně v logickém slova smyslu, to jest pro každý empirický výrok). Užití tohoto výroku tak bude s příslušným informačním stavem neslučitelné - povede ke stavu inkonsistence. (V praxi asi vyvolá nějaký ‘backtracking’ prověřující, zda byla všechna dosud přijatá tvrzení opodstatněná, ústící v nějaké rozhodnutí v tom smyslu, zda je třeba odmítnout ten poslední výrok, či už nějaký předchozí.)

Tento model nám ovšem otevírá prostor pro ten druh analýzy modalit, který potřebujeme, abychom explikovali intuici, o které jsme mluvili výše. Takto analyzovány, fungují modality jako *testy*: informační stav fakticky nemění, jenom ho buďto ‘schvalují’ nebo ‘neschvalují’. Z formálního hlediska to znamená, že každé množině možných světů přiřadí buďto tuto množinu samotnou (‘schválení’), nebo množinu prázdnou (‘neschválení’). Taková ‘epistemická’ možnost tedy bude definována následujícím způsobem:

$$\|\Diamond V\|^{\mathbf{P}}(\mathrm{M}) \quad = \mathrm{M}, \text{ jestliže } \|V\|^{\mathbf{P}}(\mathrm{M}) \neq \varnothing \text{ (tj. } \mathrm{M} \cap \|V\|^{\mathbf{I}} \neq \varnothing)$$
$$= \varnothing, \text{ jestliže } \|V\|^{\mathbf{P}}(\mathrm{M}) = \varnothing \text{ (tj. } \mathrm{M} \cap \|V\|^{\mathbf{I}} = \varnothing)$$

K tomu, abychom analyzovali (3) a (4) nyní ještě potřebujeme operátor, který zachycuje ‘zřetězení’ výroků (s obyčejnou konjunkcí zřejmě nevystačíme, protože nemají-li (3) a (4) vyjít jako ekvivalentní, musí být tento operátor nesymetrický). Takový operátor je ale v daném rámci snadné definovat:

$$\|V_1; V_2\|^{\mathbf{P}}(\mathrm{M}) = \|V_2\|^{\mathbf{P}}(\|V_1\|^{\mathbf{P}}(\mathrm{M}))$$

Rozdělme nyní danou množinu M možných světů na množinu světů, ve kterých prší (M′) množinu těch, ve kterých neprší (M″). Předpokládejme, že M′ i M″ jsou neprázdné. Pak zřejmě platí

$$\|\Diamond \textit{Prší}; \neg \textit{Prší}\|^{\mathrm{P}}(\mathrm{M}) = \|\neg \textit{Prší}\|^{\mathrm{P}}(\|\Diamond \textit{Prší}\|^{\mathrm{P}}(\mathrm{M})) = \|\neg \textit{Prší}\|^{\mathrm{P}}(\mathrm{M}) = \mathrm{M}''$$
$$\|\neg \textit{Prší}; \Diamond \textit{Prší}\|^{\mathrm{P}}(\mathrm{M}) = \|\Diamond \textit{Prší}\|^{\mathrm{P}}(\|\neg \textit{Prší}\|^{\mathrm{P}}(\mathrm{M})) = \|\Diamond \textit{Prší}\|^{\mathrm{P}}(\mathrm{M}'') = \emptyset$$

Zatímco v druhém případě je výsledkem stav 'informační frustrace', v prvním tomu tak není.

Podívejme se nyní na sémantiku výroků z velice abstraktního hlediska. Výroky přirozeného jazyka jsou pravdivé či nepravdivé, mnohé z nich ovšem v závislosti na něčem mimojazykovém; a my proto pracujeme s různými množinami 'indexů', vzhledem ke kterým je pravdivost výroků relativní (nejčastěji se jim říká 'možné světy'). Klasická analýza modalit Saula Kripka navíc naznačila, že můžeme potřebovat i něco jako 'přechody' mezi indexy: výrok je *možný* vzhledem k indexu *i* tehdy, lze-li z tohoto indexu 'přejít' k indexu, ve kterém je tento výrok pravdivý[2]. Výroky jsou přitom charakterizovány tím, vzhledem ke kterým indexům jsou pravdivé – můžeme je tedy vidět jako vyjadřující příslušné množiny indexů (kterým se pak obvykle říká *propozice*).

Srovnejme nyní z tohoto hlediska přirozený jazyk s jazyky programovacími. Ukazuje se, že k jejich analýze můžeme použít obdobný pojmový rámec, i když jiným způsobem. Opět máme 'indexy' (které tentokrát odpovídají stavům počítače) a 'přechody' mezi nimi - jenomže příkaz programovacího jazyka nyní není, na rozdíl od výroku přirozeného jazyka, charakterizován tím, vzhledem ke kterým indexům je pravdivý, ale tím, jaký typ přechodu mezi indexy vyjadřuje. Příkazy a z nich složené programy programovacích jazyků jsou tedy nahlédnutelné jako vyjadřující přechody, a nikoli propozice[3].

I v rámci programovacích jazyků ovšem najdeme výrazy, které je třeba nahlížet jako vyjadřující propozice a nikoli přechody: to jsou výrazy obvykle nazývané 'booleovské' nebo 'logické', které se používají v rámci podmíněných příkazů, cyklů apod. Z druhé strany, jak jsme právě viděli, výroky přirozeného jazyka může být z určitého hlediska užitečné nahlížet jako vyjadřující nikoli propozice, ale přechody. To naznačuje, že jak pro analýzu přirozeného jazyka, tak pro analýzu jazyků programovacích můžeme potřebovat oba tyto typy sémantických entit – rozdíl je v tom, který z nich je pro daný typ jazyka klíčový.

Logický systém, který vzniká artikulací takovýchto úvah, je tzv. *výroková dynamická logika* (*propositional dynamic logic - PDL*) rozebíraná van Benthemem (1997)[4]. V rámci této logiky máme dvě základní kategorie extralogických konstant, 'výroky' a 'programy'. Každý z nich si s sebou nese svůj typ operátorů: výroky lze negovat, konjugovat ap.; zatímco programy lze zřetězovat, iterovat atd. Sémantika této logiky je založena množině S ('stavů' či 'světů') a interpretační funkci I, přiřazující každému elementárnímu výroku podmnožinu S a každému elementárnímu programu binární relaci na S.

Sémantika výroků PDL je definována očekávaným způsobem. Klasická definice prostřednictvím splňování (které je ovšem nyní relativizováno nikoli jenom k interpretaci konstant, ale i ke 'stavům') vypadá následovně:

$\mathrm{I,s} \models V$ jestliže $\mathrm{s} \in \mathrm{I}(V)$, je-li V atomický
$\mathrm{I,s} \models \neg V$ jestliže neplatí $\mathrm{I,s} \models V$
$\mathrm{I,s} \models V_1 \wedge V_2$ jestliže $\mathrm{I,s} \models V_1$ a $\mathrm{I,s} \models V_2$
atd.

Pro programy pak platí

$\mathrm{I,s,s}' \models \pi$ jestliže $\langle \mathrm{s,s}' \rangle \in \mathrm{I}(\pi)$, je-li π atomický
$\mathrm{I,s,s}' \models \pi_1;\pi_2$ jestliže existuje s'' tak, že $\mathrm{I,s,s}'' \models \pi_1$ a $\mathrm{I,s}'',\mathrm{s}' \models \pi_2$

I,s,s′ $\models$ π^* jestliže existují $s_1,...,s_n$ tak, že $s=s_1$, $s'=s_n$ a I, s_i,s_{i+1} $\models$ π pro $i=1,...,n-1$
atd.

(Samozřejmě, že totéž bychom mohli vyjádřit i jako definice denotátů, množin ‘stavů’ přiřazovaných výrokům a dvojic stavů programům: $\|V\| = \mathrm{I}(V)$, je-li V atomický; $\|\neg V\| = S\backslash\|V\|$ $(= \{x\in S : x\notin \|V\|\})$; $\|V_1\wedge V_2\| = \|V_1\|\cap\|V_2\|$; $\|\pi\| = \mathrm{I}(\pi)$, je-li π atomický; $\|\pi_1;\pi_2\| = \{<s,s'> : \text{existuje } s'' \text{ tak, že } <s,s''>\in\|\pi_1\| \text{ a } <s'',s'>\in\|\pi_2\|\}$; $\|\pi^*\| = \{<s,s'> : \text{existují } s_1,...,s_n$ tak, že $s=s_1$, $s'=s_n$ a $<s_i,s_{i+1}>\in\|\pi\|$ pro $i=1,...,n-1\}$ atd.)

Kromě toho ovšem existují prostředky ‚interakce‘ mezi výroky a programy: od každého programu π můžeme ‘kripkovským způsobem’ odvodit příslušnou modalitu, to jest příslušný operátor π-možnosti $<\pi>$ (a samozřejmě i duální operátor π-nutnosti $[\pi]$):

I,s $\models$ $<\pi>V$ jestliže existuje s′ tak, že I,s,s′ $\models$ π a I,s′ $\models$ V

Výrok $<\pi>V$ tedy říká: programem π se lze dostat do stavu, ve kterém platí V.

Naopak na základě každého výroku V můžeme vytvořit ‘testovací‘ program $(V)?$, který dělá pouze to, že se posouvá od daného stavu k témuž stavu, je-li V v tomto stavu pravdivý, a v opačném případě končí:

I,s,s′ $\models$ $(V)?$ jestliže s = s′ a I,s $\models$ V

Různé výrokové logiky nyní můžeme vidět jako speciální případy PDL. Tak klasický výrokový počet bude PDL bez programů. Klasická modální logika bude variantou PDL s jediným programem, indukujícím standardní modality (například v případě logiky S5 by to byl program denotující ekvivalenci na S). Multimodální logiky (to jest logiky s více druhy modalit) by pak měly programů více, ale opět jenom jako induktory modalit. Při použití PDL na analýzu programovacího jazyka, jako je třeba PASCAL, by programy odpovídaly příkazům a jejich zřetězením, zatímco výroky by odpovídaly booleovským výrazům. A při analýze dynamických aspektů přirozeného jazyka by potom některým či všem větám odpovídaly nikoli výroky, ale programům[5].

Chceme-li ovšem výroky přirozeného jazyka explikovat jako to, čemu se v rámci DPL říká programy (to jest explikovat jejich významy nikoli jako propozice, množiny stavů, ale jako ‘přechody’ či PZIS, to jest binární relace mezi stavy), musíme se vypořádat ještě s jedním zásadním problémem, totiž s tím, jak takto definovaná sémantika zakládá pravdivostní podmínky a vyplývání (což je pro logiku to podstatné). Víme, že v případě intenzionální sémantiky jsou pravdivostní podmínky přímo explikovány denotáty (denotátem výroku je množina právě všech těch možných světů, v nichž je tento výrok pravdivý) a vyplývání je pak přímočaře definovatelné jejich prostřednictvím (výrok V vyplývá z výroku V' je-li denotát V' obsažen v denotátu V'). Bereme-li však za denotáty výroků binární relace mezi stavy či světy, toto propojení se zdá ztrácet.

Standardní způsob jeho reinstalace tohoto propojení je následující: výrok V je definován jako pravdivý ve stavu s právě tehdy, když existuje stav s' tak, že $<s,s'>\in\|V\|$. (Kdybychom se na V dívali jako na program, mohli bychom říci, že pravdivost vzhledem ke stavu s je definována jako schopnost proběhnout s s jako počátečním stavem.) Intuice za touto definicí je následující: výrok, jako je *Prší*, je v daném stavu pravdivý právě tehdy, když je s tímto stavem slučitelný, to jest když není v rozporu s fakty danými tímto stavem; pak ovšem tento výrok vede od tohoto stavu ke stavu jinému, v prototypickém případě informačně bohatšímu.

Jak se nyní od takovéto dynamické výrokové logiky posunout k logice predikátové? Mohli bychom to samozřejmě učinit zcela triviálně. Je ovšem také možné vzít v úvahu fakt, že podstatné 'dynamické' aspekty sémantiky přirozeného jazyka vznikají právě na úrovni přístupné až predikátové logice, a pokusit se o jejich explikaci.

Vraťme se k problematice anglického neurčitého členu, kterou jsme se zabývali v předchozím pokračování tohoto seriálu. Vezměme anglickou verzi věty *Nějaký člověk jde*, tedy

(5) A man walks.

To, co je od Russellových dob přijímáno za její standardní (extenzionální) analýzu, totiž

(5′) $\exists x\,(\mathbf{man}(x) \wedge \mathbf{walk}(x))$,

je, jak jsme viděli, ne zcela adekvátní proto, že (5) normálně nejenom konstatuje existenci jdoucího muže, ale navíc tohoto muže 'uvádí na scénu' (= nějakým způsobem ho činí součástí kontextu, který produkuje), takže k němu pak může odkazovat nějaká další věta, třeba *The man whistles* [*Ten člověk si píská*]. A tohle je něco, co se nezdá být nijak zachytitelné prostředky, které nám poskytuje standardní (ať už extenzionální či intenzionální) logika.

Jednou s cest, jak právě tuhle intuici učinit východiskem dynamické predikátové logiky, je ta, kterou se vydali Groenendijk a Stokhof (1991) se svou *dynamickou predikátovou logikou* (*DPL*) a o které jsem již v Organonu psal (viz Peregrin, 1996). Zatímco v rámci standardní logiky je formule, jako je $\mathbf{man}(x)$, pravdivá relativně k odhonocení proměnných (konkrétně $\mathbf{man}(x)$ je pravdivá vzhledem k těm a jenom těm ohodnocením, které proměnné x přiřazují muže), v rámci DPL je pravdivost formulí relativizována vzhledem k *dvojicím* ohodnocení proměnných. Prvky takové dvojice se přitom chápou jako ohodnocení 'vstupní' a 'výstupní' – formule je tedy, můžeme říci, chápána jako něco, co je vyhodnocováno s jistým ohodnocením 'na vstupu' a z čeho jiné ohodnocení vychází 'na výstupu'.

Přitom u většiny formulí se 'výstupní' ohodnocení neliší od 'vstupního' – to jest tyto formule jsou pravdivé jenom vzhledem k takovým dvojicím ohodnocení, jejichž obě složky jsou identické. Takže například sémantické pravidlo pro atomické výroky je v podstatě jenom triviální variantou pravidla známého ze standardního predikátového počtu (pro jednoduchost bez funkčních symbolů):

$\mathrm{I,v,v'} \models P(t_1,...,t_n)$ právě tehdy, když $\mathrm{v}=\mathrm{v}'$ a $<\|t_1\|,...,\|t_n\|>\in \mathrm{I}(P)$, kde $\|t\|=\mathrm{I}(t)$, je-li t konstanta a $\|t\|= \mathrm{v}(t)$, je-li t proměnná

Existují ovšem i výroky, u kterých tomu tak není, to jest které mohou produkovat 'výstupní' ohodnocení lišící se od ohodnocení 'vstupního'. V případě DPL to jsou především výroky s existenčním kvantifikátorem:

$\mathrm{I,v,v'} \models \exists xV$ právě tehdy, když existuje v'' tak, že $\mathrm{I,v'',v'} \models V$ a v'' se od v liší nanejvýše v hodnotě přiřazované proměnné x.

Jediným dalším typem výroku, u nějž se ‘výstupní’ ohodnocení nemusí shodovat s ohodnocením ‘vstupním’, je konjunkce, která je definována analogicky výše probíranému ‘zřetězení‘

$I,v,v' \models V_1 \wedge V_2$ právě tehdy, když existuje v'' tak, že $I,v,v'' \models V_1$ a $I,v'',v' \models V_2$,

(Vidíme ovšem, že konjunkce sama žádné dynamické efekty neprodukuje, jenom je případně promítá ze spojovaných výroků na jejich spojení. Jediným typem operátoru, který dynamický efekt skutečně produkuje, tak v rámci PDL zůstává existenční kvantifikátor.)

Situaci tedy můžeme vidět následujícím způsobem: zatímco formulí $\mathbf{man}(x)$ ohodnocení beze změny buď prostě ‘projde’, nebo ‘neprojde’ (podle toho, přiřazuje-li proměnné x muže, nebo ne), v případě formule $\exists x\ \mathbf{man}(x)$ se ohodnocení nejprve změní tak, aby matricí této formule, kterou je opět $\mathbf{man}(x)$, pokud možno prošlo (to jest změní se hodnota, kterou toto ohdonocení přiřazuje proměnné x tak, aby to byl muž – pokud ovšem nějaký muž v univerzu existuje).

Formule $\exists x\ \mathbf{man}(x)$ je tedy, tak jako v klasickém případě, splnitelná (a tudíž, protože je uzavřená, i pravdivá) právě tehdy, když v univerzu existuje alespoň jeden muž; avšak navíc ‘nastavuje’ hodnotu proměnné x (na muže). To znamená, že formule DPL

(6) $(\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$

je splnitelná za týchž podmínek, za nichž je splnitelná formule (5′) standardní logiky – to jest právě tehdy, když v univerzu existuje muž, který jde. To lze ukázat následujícím způsobem. Podle definice je tomu tak, že

$I,v,v' \models (\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ právě tehdy, když existuje v'' tak, že $I,v,v'' \models \exists x\ \mathbf{man}(x)$ a $I,v'',v' \models \mathbf{walk}(x)$.

Podle definice dále platí, že $I,v,v'' \models \exists x\ \mathbf{man}(x)$ právě tehdy, když existuje v''' tak, že $I,v''',v'' \models \mathbf{man}(x)$ a v''' se od v liší nanejvýše v hodnotě přiřazované proměnné x; a $I,v'',v' \models \mathbf{walk}(x)$ právě tehdy, když $v' = v''$ a $v''(x) \in I(\mathbf{walk})$. Dostáváme tedy

$I,v,v' \models (\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ právě tehdy, když existují v'' a v''' tak, že $I,v''',v'' \models \mathbf{man}(x)$, v''' se od v liší nanejvýše v hodnotě přiřazované proměnné x, a $v' = v''$ a $v''(x) \in I(\mathbf{walk})$.

Zřejmými úpravami se dále dostáváme k vyjádření

$I,v,v' \models (\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ právě tehdy, když existuje v''' tak, že $I,v''',v' \models \mathbf{man}(x)$, v''' se od v liší nanejvýše v hodnotě přiřazované proměnné x a $v'(x) \in I(\mathbf{walk})$.

A protože podle definice dále platí, že $I,v''',v' \models \mathbf{man}(x)$ právě tehdy, když $v''' = v'$ a $v'''(x) \in I(\mathbf{man})$, dostáváme

$I,v,v' \models (\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ právě tehdy, když existuje v''' tak, že $v''' = v'$, $v'''(x) \in I(\mathbf{man})$, v''' se od v liší nanejvýše v hodnotě přiřazované proměnné x a $v'(x) \in I(\mathbf{walk})$, a po dalších úpravách

I,v,v′ $\models$ $(\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ právě tehdy, když se v′ od v liší nanejvýše v hodnotě přiřazované proměnné x a $v'(x) \in I(\mathbf{walk}) \cap I(\mathbf{man})$.

To tedy znamená, že formule $(\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x)$ je splnitelná právě tehdy, když mají extenze I(**walk**) a I(**man**) neprázdný průnik.

To ukazuje, že v rámci DPL existenční kvantifikátor vlastně váže proměnnou i napravo od toho, co by bylo v rámci standardní logiky jeho dosahem. Tento podivný fakt se poněkud objasní, uvědomíme-li si, že tento kvantifikátor *de facto* spíše než jako kvantifikátor v tradičním slova smyslu funguje jako přiřazovací příkaz programovacího jazyka – přiřadí proměnné hodnotu, kterou si tato podrží až do té doby, než jí je přiřazena nějaká jiná. (Přitom však, jak jsme viděli, i tento dynamický kvantifikátor produkuje pravdivostní podmínky, které charakterizují kvantifikátor standardní!) Proměnné DPL tak ovšem přestávají být proměnnými v tradičním slova smyslu (a proto o nich autoři DPL také hovoří jako o *diskurzních značkách*).

V rámci DPL tedy existenční kvantifikátor funguje jako něco, co pomáhá uvést 'na scénu' individuum, ke kterému je pak možné dále odkazovat; a tak může sloužit jako příměřenější nástroj analýzy výroků s anglickým neurčitým členem. Z hlediska obecné teorie dynamické logiky je DPL speciálním případem, ve kterém je informační stav *de facto* ztotožněn s ohodnocením proměnných. (Jde ovšem o systém, který zůstává na čistě extenzionální úrovni – bylo by však jistě možné uvažovat i o intenzionálních variantách[6].)

V kontextu DPL se do jisté míry můžeme vypořádat i s určitým členem. V minulém pokračování jsme viděli, že s jeho 'post-russellovskou' analýzou (která za jeho denotát bere funkci přiřazující jednoprvkové extenzi její jediný prvek a každé jiné nic) nevystačíme – tato analýza totiž zjevně vede k závěru, že věta jako *The man whistles* [*Ten muž si hvízdá*] nemůže být (v důsledku toho, že jistě existuje více než jediný muž) pravdivá. Intuice nám však říká, že ta věta docela dobře pravdivá být může - protože fráze *the man* běžně odkazuje nikoli k jedinému muži na světě, ale k *tomu jedinému muži, o němž byla dosud řeč*. V rámci DPL můžeme tuto větu (stejně tak jako větu *On si hvízdá*) analyzovat prostě jako **whistle**(x). Z toho, co jsme řekli výše, totiž vyplývá, že formule, která vznikne, když **whistle**(x) prostě pomocí konjunkce připojíme za analýzu věty *A man walks*, totiž formule

(7) $(\exists x\ \mathbf{man}(x)) \wedge \mathbf{walk}(x) \wedge \mathbf{whistle}(x)$,

bude mít, jakožto formule DPL, přesně ty pravdivostní podmínky, které připisujeme větě *A man walks and the man* [*he*] *whistles*, to jest podmínky vyjadřované standardním, nedynamickým výrokem

(7′) $\exists x\ (\mathbf{man}(x) \wedge \mathbf{walk}(x) \wedge \mathbf{whistle}(x))$.

Z hlediska analýzy přirozeného jazyka lze ovšem Groenendijkovo a Stokhofovo řešení považovat za poněkud *ad hoc*. (To je patrné i z toho, že právě uvedená analýza funguje jenom díky tomu, že jsme při analýze věty *The man whistles* 'náhodou' zvolili právě diskurzní značku x – je zřejmé, že kdybychom byli zvolili třeba y, výsledek by byl méně uspokojivý.) Mechanismus 'anaforické reference' v přirozeném jazyce totiž funguje na poněkud jiných principech. Nejnázorněji je to vidět na angličtině s jejími členy: *an F* v typickém případě uvádí ‘na scénu’ nějakého reprezentanta objektů splňujících *F* a *the F* se pak právě na tohoto reprezentanta odvolává. Kdybychom tedy chtěli logiku, která tento mechanismus zachycuje přímo, museli bychom zavést nějaký aparát, který by nám dovoloval ustanovovat a měnit reprezentanty různých pojmů (či v extenzionální variantě extenzí těchto pojmů) a museli bychom tedy předpokládat, že máme nějakou (ne nutně totální) funkci, přiřazující pojmům (či podmnožinám univerza) jejich reprezentanty.

Jednou takovou funkcí je ta, která vyplývá z russellovské analýzy určitého členu (viz předchozí pokračování tohoto seriálu): funkce, která každé jednoprvkové množině přiřadí její jediný prvek a pro žádnou jinou množinu není definována. Jak jsme ovšem viděli, tato analýza není obecně přijatelná. V kontextu dynamické logiky ovšem máme jinou možnost: uchopit výběrovou funkci jako součást kontextu, to jest jako něco, co se v průběhu vyhodnocování mění. Reprezentant pojmu *P*, ke kterému odkazuje výraz **the** *P*, tedy může být dynamicky nastavován třeba užitím výrazu **a** *P*.

Tyto intuice lze formalizovat například následujícím způsobem: Nazvěme *výběrovou funkcí* parciální funkci na množině všech podmnožin univerza, která každé množině ve svém definičním oboru přiřazuje prvek této množiny. Předpokládejme, že predikáty jsou interpretovány jako v rámci standardní (extenzionální) logiky (tj. jako označující množiny individuí) a definujme interpretaci výrazů tvaru **a** *P* a **the** *P* tak, aby označovaly množiny dvojic výběrových funkcí:

$$\|\mathbf{a}\ P\| = \{<c,c'> \mid c(s) = c'(s) \text{ pro každou } s \neq \|P\| \text{ a } c'(\|P\|) \in \|P\|\}$$
$$\|\mathbf{the}\ P\| = \{<c,c> \mid c(\|P\|) \in \|P\|\}$$

Definujme dále *referent* výrazů **a** *P* a **the** *P* vzhledem k výběrové funkci c následujícím způsobem:

$$|\mathbf{a}\ P|_c = |\mathbf{the}\ P|_c = c(\|P\|)$$

Pak můžeme definovat interpretaci výroků předpisem

$$\|P(T)\| = \{<c,c'> \mid <c,c'> \in \|T\| \text{ a } |T|_{c'} \in \|p\|\}$$
$$\|V_1 \wedge V_2\| = \{<c,c'> \mid \text{existuje } c'' \text{ tak, že } <c,c''> \in \|V_1\| \text{ a } <c'',c'> \in \|V_2\|\}$$

Pak lze například snadno ukázat, že P_1(**a** *P*)$\wedge P_2$(**the** *P*) je pravdivý právě když existuje prvek univerza, který má vlastnosti *P*, P_1 a P_2 – to znamená, že například výrok **walk** (**a man**) $\wedge$ **whistle** (**the man**) je pravdivý právě tehdy, když je v univerzu přítomno individuum, které patří do extenzí všech tří predikátů **man**, **walk** a **whistle** - to jest je to muž, který jde a hvízdá si. (Podrobněji viz Peregrin & von Heusinger, 1997, a Peregrin, 2000).

Nyní se můžeme vrátit k tomu, co jsme řekli v předchozím pokračování o neurčitém členu: souvisí-li jeho funkce s uváděním objektů ‘na scénu’, pak jeho význam nemůžeme explikovat jinak než v rámci dynamické sémantiky, totiž jako přechod od daného kontextu ke kontextu s nově zavedeným objektem. Podobně je to pak i s určitým členem, jehož častou funkcí je odkazovat k takto ‘kontextuálně vysvíceným’ objektům.

Přitom je dobré si uvědomit, že takovouto ‘kinematiku diskurzu’ nelze z hlediska logiky prostě pominout jako věc pragmatiky – zjevně tu totiž jde mimo jiné i o určité *instance vyplývání* (z čistě logického hlediska možná ne příliš zajímavé, nicméně v přirozeném jazyce nepochybně existující): takové instance, jako je ekvivalence výroků "Nějaký člověk jde a hvízdá si" a "Nějaký člověk jde a ten člověk si hvízdá". Z tohoto hlediska je tedy problematické i to, když je přechod od standardní logice k logice dynamické chápán jako *de facto* posun od logiky v pravém slova smyslu (nauky o vyplývání, to jest o tom, jak pravdivost některých výroků zaručuje pravdivost jiných výroků) k epistemologii či kognitivní psychologii (k teorii toho, jak lidé usuzují či chápou)[7].

POZNÁMKY:

1. V lingvisticky orientovaných teoriích se obvykle hovoří o *pontenciálu změny kontextu*.

2. Tato potřeba je naprosto zřejmá například v případě modální analýzy časů přirozeného jazyka. Řeknu-li, že bude pršet, pak tím říkám nejenom to, že existuje nějaký časový okamžik, ve kterém prší, ale navíc i to, že jde o okamžik, *který je v určitém vztahu k okamžiku aktuálnímu* (totiž následuje jej na časové ose).

3. Viz např. Gordon (1979).

4. Klasickými texty jsou v tomto směru Harel (1984) či Goldblatt (1987).

5. Výše diskutovanému operátoru 'epistemické možnosti' by tak odpovídal modální operátor aplikovatelný na programy: $I,s,s' \models \Diamond\pi$ právě tehdy, když $s = s'$ a existuje s'' tak, že $I,s,s'' \models \pi$.

6. Autoři DPL publikovali i výsledky týkající se přenesení myšlenek DPL na Montaguovu intenzionální logiku – viz Groenendijk a Stokhof (1990).

7. Van Benthem (1997, s. ix) například říká, že dynamická logika zachycuje „logickou strukturu kognitivních akcí, stojících v základě lidského souzení nebo chápání přirozeného jazyka".

LITERATURA:

[1] BENTHEM, J. VAN (1997): **Exploring Logical Dynamics**, CSLI, Stanford.
[2] GOLDBLATT, R. (1987): **Logics of Time and Computation**, CSLI, Stanford.
[3] GORDON, M.J.C. (1979): **The Denotational Description of Programming Languages**. Springer, New York.
[4] GROENENDIJK, J. & STOKHOF, M. (1989): Dynamic Montague grammar. **Papers from the Second Symposium on Logic and Language**, ed. L. Kálmán a L.Pólos, Akadémiai Kiadó, Budapest.
[5] GROENENDIJK, J. & STOKHOF, M. (1991): Dynamic Predicate Logic. **Linguistics and Philosophy** 14, 39-101.
[6] HAREL, D. (1984): Dynamic Logic. **Handbook of Philosophical Logic II**, ed. D.M. Gabbay & F. Guenthner, Reidel, Dordrecht, 497-604.
[7] PEREGRIN, J. (1996): Dynamická sémantika. **ORGANON F** 4, 333-348.
[8] PEREGRIN, J. (2000): The Logic of Anaphora. **Logica Yearbook 1999**, to appear.
[9] PEREGRIN, J. & VON HEUSINGER, K. (1997): Dynamic Semantics with Choice Functions. **Context-Dependence in the Analysis of Linguistic Meaning** I, ed. H. Kamp a B.Partee, Universität Stuttgart, Stuttgart, 329-354; vyjde jako kniha v nakladatelství Elsevier.

IV.
Lambekovy kalkuly a 'logiky kategorií'

Fregovým geniálním nápadem, ze kterého vyšel při svém pokusu o vytvoření rigorózní, ne-psychologické teorie významu, bylo rekonstruovat význam složeného výrazu principiálně jako výsledek funkční aplikace významu jedné z jeho částí na význam části nebo částí zbývajících. Tak význam výroku skládajícího se z podmětu a predikátu nahlédl jako výsledek aplikace významu přísudku na význam podmětu; zatímco například význam složeného výroku vytvořeného spojením dvou výroků pomocí logické spojky jako výsledek aplikace významu této spojky na významy oněch dvou výroků. Významy některých výrazů tedy stotožnil s funkcemi v matematickém slova smyslu.

Tato myšlenka došla zobecnění v rámci toho, čemu se dnes v lingvistice říká *kategoriální gramatika* (viz např. Oeherle et al., 1988), a co *de facto* stojí také v základě jazyka Churchovy (1940) 'jednoduché teorie typů' a potažmo například i jazyka TIL[1]. V rámci kategoriální gramatiky jsou výrazy rozděleny do kategorií, které jsou označeny tak, že každé gramatické pravidlo kombinuje výraz kategorie $A/A_1,...,A_n$ s výrazy kategorií A_1, ..., A_n ve výraz kategorie A. Jazyky tohoto typu se pak zcela přirozeně pojí s fregovskou, funkcionální sémantikou: významem výrazu kategorie $A/A_1,...,A_n$ je funkce, která významům výrazů kategorií $A_1,...,A_n$ přiřadí význam příslušného výrazu kategorie A. Označíme-li kategorii termů **T** a kategorii výroků **V**, můžeme tímto způsobem kategorii unárních predikátů označit **V**/**T** (a obecněji *n*-árních predikátů **V**/**T**,...,**T**) – neboť predikát je výraz, který spolu s výrazem kategorie **T** dává výraz kategorie **V**. Významem predikátu je pak funkce, která významu výrazu kategorie **T** přiřadí význam výrazu kategorie **V**[2]. (Budeme-li tedy například považovat za významy výrazů kategorie **V** pravdivostní hodnoty a za významy výrazů kategorie **T** individua, bude významem výrazu kategorie **V**/**T** funkce přiřazující pravdivostní hodnoty individuím.) Podobně kategorii větných spojek, jejíž výrazy se spojují s dvojicemi výroků ve výrok, můžeme označit **V**/**V**,**V**[3].

Vezměme větu *Někdo jde a někdo zpívá*. V rámci kategoriální gramatiky bývá obvykle analyzována následujícím způsobem:

```
 Někdo    jde     a     někdo  zpívá
 V/(V/T) V/T  V/V,V  V/(V/T)  V/T
└──────────┘        └─────────────┘
      V                    V
└───────────────────────────────────┘
                 V
```

To znamená, že tato věta je nahlížena jako výsledek spojejí spojky *a* (kategorie **V**/**V**,**V**) s dvojicí výroků (kategorie **V**), z nichž každý je výsledkem spojení 'kvantifikátoru' (výrazu kategorie **V**/(**V**/**T**)) s predikátem (výrazem kategorie (**V**/**T**)).

Vezměme ale nyní větu *Někdo jde a zpívá* (která, jak je zřejmé, není s tou předchozí synonymní) – jak bychom analyzovali ji? Zjevně bychom potřebovali, aby její přísudek vyšel jako výraz kategorie **V**/**T** – avšak tuto kategorii očividně nemůžeme kombinací dvou výrazů kategorie **V**/**T** (*jde* a *zpívá*) a jednoho výrazu kategorie **V**/**V**,**V** (*a*) dostat. Co víc, tyto tři výrazy nelze v rámci kategoriální gramatiky skombinovat vůbec nijak.

```
 Někdo    jde      a       zpívá
V/(V/T)  V/T   V/V,V      V/T
         └───────────────────┘
                   ?
```

Zdá se, že jediný způsob, jak dospět k rozumné analýze, je připustit možnost konjunktivního spojování nikoli jen vět, ale i predikátů. Na to už je ale kategoriální gramatika, zdá se, krátká: jsou-li predikáty, jako *jde* či *zpívá*, kategorie **V**/**T** a je-li *a* kategorie **V**/**V**,**V**, pak v rámci kategoriální gramatiky prostě není způsob, jak je skombinovat.

Jak tento problém řeší Churchova 'teorie typů' a potažmo TIL i jiné systémy, je známo: ke kategoriální gramatice se přidávají proměnné a mechanismus lambda-abstrakce. Predikáty P_1 a P_2 kategorie **V**/**T** se pak mohou formálně aplikovat na proměnnou x, čímž vzniknou výroky-nevýroky $P_1(x)$ a $P_2(x)$ – po věcné stránce to sice žádné výroky nejsou, protože nemají skutečný podmět, ale po stránce formální to jsou výrazy kategorie **V**, a tudíž jsou skombinovatelné se spojkou *a* (či jejím standardním logickým protipólem $\wedge$). Tak vznikne 'výrok' $P_1(x)\wedge P_2(x)$, který pak ovšem může být zbaven svého podmětu-nepodmětu x za vzniku predikátu (tj. výrazu kategorie **V**/**T**) $\lambda x.P_1(x)\wedge P_2(x)$. Větu *Někdo jde a zpívá* bychom pak mohli analyzovat následujícím způsobem[4]:

```
 Někdo   λx jde x    a    zpívá x
V/(V/T)     V/T T  V/V,V   V/T  T
            └───┘          └────┘
              V              V
            └───────────────────┘
                      V
        └───────────────────────┘
                 V/T
└───────────────────────────────┘
              V
```

Může nás ovšem napadnout, proč celou věc vyřizovat tak komplikovaně, když by to šlo jednodušeji. Proč prostě kategoriální gramatiku nerozšířit o nové typy pravidel, kterým by na úrovni sémantiky odpovídala ne aplikace funkce na argument, ale skládání funkcí? Proč přímo nepovolit kombinování výrazu kategorie A/B s výrazem kategorie B/C ve výraz kategorie A/C – když sémanticky je toto pravidlo zcela přirozené, jakožto vyjádření složení funkcí vyjadřovaných skládanými výrazy? Tak například vyjadřuje-li negace (výraz kategorie **V**/**V**) známou funkci N (přiřazující pravdivostním hodnotám pravdivostní hodnoty) a vyjadřuje-li nějaký predikát (výraz kategorie **V**/**T**) funkci F z individuí do pravdivostních hodnot, proč by je nemělo být možné přímo skombinovat v negativní predikát, vyjadřující funkci G takovou, že pro každý přípustný argument a platí $G(a) = N(F(a))$? (Bude-li tedy F například funkcí přiřazující *pravdu* právě všem vousatým logikům, bude G funkce, která přiřadí individuu *pravdu* právě tehdy, není-li to vousatý logik.) Je-li V_1 výraz kategorie A/B a V_2 výraz kategorie B/C, mohli bychom tento jejich nový druh spojení ve výraz kategorie A/C zapisovat třeba jako $\{V_1\ V_2\}$. (Z toho, co jsme právě konstatovali, vyplývá, že do lambda-kalkulu by pak $\{V_1\ V_2\}$ bylo 'přeložitelné' jako $\lambda x.V_1(V_2(x))$, kde x je proměnnou kategorie C.)

Tím se dostáváme k zajímavým úvahám o tom, jaké typy pravidel by bylo možné ke kategoriální gramatice smysluplně přidat. Než v nich však budeme pokračovat, všimněme si, že stejného účinku jako přijetím právě uvedeného nového pravidla můžeme dosáhnout i tím, že stanovíme, že na každý výraz V kategorie A/B se současně lze dívat i jako na určitý výraz kategorie $(A/C)/(B/C)$, totiž jako na ten výraz, který bychom v rámci lambda-kalkulu zapsali jako $\lambda y.\lambda x.\mathrm{V}(y(x))$ (kde x je proměnná kategorie C a y proměnná kategorie B/C). Tak například chceme-li dosáhnout skombinovatelnosti negace s predikátem v predikát, můžeme namísto ustanovení nového pravidla umožňujícího skombinovat výraz kategorie **V/V** s výrazem kategorie **V/T** ve výraz kategorie **V/T**, ustanovit, že negace se kromě jako výraz kategorie **V/V** může chovat i jako výraz kategorie (**V/T**)/(**V/T**). Takže formální alternativou přidání nových pravidel toho typu, jaké jsme zavedli výše, je přidání možnosti určitého druhu fluktuace výrazů mezi kategoriemi – například ztotožnění V kategorie A/B s určitým výrazem kategorie $(A/C)/(B/C)$.

Jak ale můžeme nechat výraz fluktuovat mezi kategoriemi? Kategorie výrazu přece určuje typ objektu, který je tímto výrazem denotován, a přesuneme-li ho do jiné kategorie, musíme mu přidělit i jiný denotát a tedy z něj *de facto* udělat zcela jiný výraz. Odpověď je v tom, že mezi objekty různých sémantických domén existují určité 'přirozené projekce': někdy se stává, že objekty jedné domény mají jisté přirozené 'protipóly' v jiné doméně, a že je pro některé účely můžeme s těmito protipóly ztotožňovat a tak dosáhnout toho, že jedna doména může být jakoby 'vnořena' do jiné. Příkladem může být 'vnoření' kategorie **T** do kategorie **V/T** založené na ztotožňování individua s jednoprvkovou množinou tvořenou právě tímto individuem. Jiným příkladem je výše popsané vnoření domény A/B do $(A/C)/(B/C)$, založené na ztotožnění funkce V s funkcí $\lambda y.\lambda x.\mathrm{V}(y(x))$ – tedy například negace $\neg$ s operací doplňku, $\lambda y.\lambda x.\neg(y(x))$.

Vraťme se nyní k obecnému problému rozšiřování kategoriální gramatiky. Samozřejmě, že nechceme-li prostě zlikvidovat fregovský 'funkcionální' model sémantiky, nemůžeme připustit skombinovatelnost *jakéhokoli* výrazu s *jakýmkoli* jiným výrazem; či vnořování *jakékoli* domény do *jakékoli* jiné. Můžeme připustit jenom takové kombinace či taková vnořování, které jsou v rámci tohoto modelu rozumně sémanticky ošetřitelné. Které to ale obecně jsou? Tuto obecnou otázkou můžeme položit následujícím způsobem: pro které kategorie $A_1, ..., A_n, A$ je rozumné připustit pravidlo kombinující výrazy kategorií $A_1, ..., A_n$ ve výraz kategorie A? (V rámci standardní kategoriální gramatiky by byla odpověď jednoduchá: pro ty, kde pro některé i platí $A_i = A/A_1,...,A_{i-1},A_{i+1},...,A_n$.). Anebo: pro které kategorie A a B je rozumné připustit, že výrazy kategorie A se mohou chovat i jako výrazy kategorie B? (Na tuto otázku by v rámci standardní kategoriální gramatiky byla odpověď, že jedině pro $B = A$.)

Učiňme nyní něco na první pohled zcela nesmyslného: podívejme se na symboly, kterými v rámci kategoriální gramatiky označujeme jednotlivé kategorie, jako na formule výrokového počtu. Představme si, že písmena v těchto symbolech jsou jednoduchými výrokovými symboly a že lomítko značí inverzní implikaci; takže například A/B budeme číst jako $B \to A$. Navíc konstatování 'výrazy kategorií $A_1, ..., A_n$ lze skombinovat ve výraz kategorie A' čtěme jako konstatování toho, že výrok A vyplývá z výroků $A_1, ..., A_n$, zkráceně $A_1,...,A_n \Rightarrow A$. To znamená, že například fakt, že v rámci kategoriální gramatiky můžeme výrazy kategorií A/B a B skombinovat ve výraz kategorie A nahlédneme jako tvrzení, že výrok A vyplývá z výroků $B \to A$ a B:

$$B \to A, B \Rightarrow B.$$

Podobně to, co by mělo platit v námi uvažovaném rozšíření kategoriální gramatiky, totiž že výrazy kategorií A/B a B/C můžeme skombinovat ve výraz kategorie A/C, nahlédneme jako tvrzení

$$B \to A, C \to B \Rightarrow C \to A.$$

Podíváme-li se nyní na to, co jsme touto transformací dostali, blížeji, zjistíme, že obě uvedená tvrzení jsou platná: dostali jsme dvě platní inferenční pravidla.

Na tom by nebylo nic zajímavého, pokud by to byla náhoda. Dá se však ukázat, že to náhoda není. Zamyslíme-li se nad celou věcí podrobněji, zjistíme, že paralela mezi otázkami *které kategorie lze skombinovat z jiných?* a *které implikativní formule vyplývají z jiných?* je zcela systematická; a že obecně platí, že výrazy kategorií A_1, ..., A_n lze v rámci lambda-kalkulu skombinovat ve výraz kategorie A právě tehdy, když je možné z A_1, ..., A_n vyvodit A v rámci intuicionistické implikativní logiky. Proč tomu tak je, není těžké nahlédnout. Uvažme totiž způsob, jak je intuicionistická implikace definována v rámci systému přirozené dedukce (viz např. Prawitz, 1965):

$$\frac{A \quad A \to B}{B} \qquad\qquad \frac{\begin{array}{c}[A]\\ \vdots \\ B\end{array}}{A \to B}$$

Tedy: (i) z A a $A \to B$ vyplývá B; a (ii) můžeme-li z předpokladu A dokázat B, pak analogicky bez předpokladu A můžeme dokázat $A \to B$. Nyní snadno nahlédneme, že ‘kombinovatelnost‘ v rámci lambda-kalkulu je definována strukturálně zcela analogicky: (i) výrazy kategorií A a B/A lze skombinovat ve výraz kategorie A; a (ii) můžeme-li za pomoci výrazu kategorie A vytvořit výraz kategorie B, pak analogicky bez použití A (za pomoci příslušné proměnné) můžeme vytvořit výraz kategorie B/A.

Každému diagramu ‘odvození’ kategorie složeného výrazu z kategorií jeho částí (to jest diagramu toho typu, jaký jsme výše uvedli například pro větu *Někdo jde a někdo zpívá*) tedy odpovídá skutečné odvození formule odpovídající kategorii složeného výrazu z formulí odpovídajících kategoriím jeho částí. Například v případě věty *Někdo jde a někdo zpívá* by tímto odvozením bylo

$$\frac{\dfrac{(\mathrm{T}\to\mathrm{V})\to\mathrm{V} \quad \mathrm{T}\to\mathrm{V}}{\mathrm{V}} \qquad (\mathrm{V}\to\mathrm{V})\to\mathrm{V} \qquad \dfrac{(\mathrm{T}\to\mathrm{V})\to\mathrm{V} \quad \mathrm{T}\to\mathrm{V}}{\mathrm{V}}}{\mathrm{V}}$$

V některých učebnicích logiky bývá uváděno, že to, s čím se setkáváme například u výrokového počtu, je čistě abstraktní struktura, jejíž je výroková logika jenom jednou možnou instancí, která však může mít instance i zcela jiné. Jako příklad takové jiné instance pak bývají téměř univerzálně uváděny elektrické obvody s ‘logickými’ prvky. Nyní vidíme mnohem neotřelejší příklad stejného jevu: zjišťujeme, že struktura, která charakterizuje logické vyplývání, je vlastní i fenoménu s vyplýváním nijak přímo nesouvisejícím.

Zjemněme nyní poněkud způsob, kterým hovoříme o kombinování výrazů. Když jsme říkali, že nějaký výraz patří do kategorie A/B, říkali jsme tím, že může být (určitým způsobem) spojen s výrazem kategorie B ve výraz kategorie C. Rozlišme teď dva typy připojení (zleva a zprava): říkejme, že výraz je kategorie A/B, platí-li, že připojí-li se k němu *zprava* výraz kategorie B, vznikne výraz kategorie A; a že je kategorie $B\backslash A$, platí-li, že připojí-li se k němu *zleva* výraz kategorie B, vznikne výraz kategorie A. Navíc říkejme, že výraz je kategorie $A \bullet B$, skládá-li se z výrazu kategorie A následovaného výrazem kategorie B.

Předpokládejme nyní, že výrazy kategorie A mají tu vlastnost, že když se k nim zprava připojí výraz kategorie B, vznikne výraz kategorie C; jinými slovy, že výrazy kategorie A jsou současně kategorie C/B. Pak je zřejmě každý výraz kategorie A následovaný výrazem kategorie B výrazem kategorie C, to jest každý výraz kategorie $A \bullet B$ je výrazem kategorie C. Navíc pro každý výraz kategorie B platí, že připojí-li se k němu zleva výraz kategorie A, vznikne výraz kategorie C, to jest každý výraz kategorie B je výrazem kategorie $A\backslash C$. To znamená

$$A \Rightarrow C/B \text{ právě když } A \bullet B \Rightarrow C \text{ právě když } B \Rightarrow A\backslash C.$$

Relace $\Rightarrow$ je navíc zjevně reflexivní a tranzitivní, tj. platí:

$$A \Rightarrow A$$
$$\text{jestliže } A \Rightarrow B \text{ a } B \Rightarrow C, \text{ pak } A \Rightarrow C$$

Tři právě uvedená tvrzení tvoří axiomy jednoho z logických systémů, jejichž prostřednictvím navrhl Lambek (1958; 1961) studovat 'matematiku větné struktury'. (Lambekovy práce původně nevzbudily téměř žádnou pozornost; uznání došly až s asi dvacetiletým zpožděním, kdy se logici a matematici pustili do systematického zkoumání kategoriálních gramatik.) Co je pozoruhodné, je to, že logiky tohoto druhu mohou být považovány za specifický druh modálních logik, totiž že je možné je sémanticky interpretovat způsobem, který je – z formálního hlediska – případem možnosvětové sémantiky navržené Kripkem (1963) pro modální logiky.

Představme si, že máme množinu W 'možných světů' a ternární relaci R 'dosažitelnosti' mezi nimi (to jest $R \subseteq W^3$). Pro výše uvedené kategoriální symboly nahlédnuté jako výroky je pak možné definovat následující sémantiku:

$$\|A \bullet B\| = \{z \mid \mathrm{R}zxy \text{ pro nějaké } x \in \|A\| \text{ a } y \in \|B\|\}$$
$$\|A/B\| = \{x \mid \text{jestliže } \mathrm{R}zxy \text{ a } y \in \|B\|, \text{ pak } z \in \|A\|\}$$
$$\|B\backslash A\| = \{x \mid \text{jestliže } \mathrm{R}zyx \text{ a } y \in \|B\|, \text{ pak } z \in \|A\|\}$$

Dá se ukázat, že výše uvedený Lambekův kalkul je vzhledem k této sémantice úplný.

To znamená, že způsoby syntaktického kombinování výrazů můžeme z tohoto pohledu vidět jako *modality* – v této souvislosti se pak hovoří o *multimodálních gramatikách*.

Úvahy naznačené v tomto článku daly vzniknout odvětví logiky, které se v poslední době bouřlivě rozvíjí (zejména v Holandsku) a pro které by bylo snad nejpřijatelnějším českým názvem *logika kategorií* (*viz* Morill, 1994, a Moortgart, 1997). Jak už jsme viděli, nejedná se vlastně o logiku ve vlastním slova smyslu, ale *o aplikaci deduktivních struktur, které logika vyvinula pro studium vyplývání, na něco, co nemá s vyplýváni nic společného, totiž na syntax, přesněji na syntaktickou kombinovatelnost* (přičemž ne všichni autoři, kteří o daném tématu píší, si tohle dostatečně jasně uvědomují.)[5]

POZNÁMKY:

[1]. Za ‘praotce’ kategoriální gramatiky bývají považováni Ajdukiewicz (1935) a Bar-Hillel (1953).

[2]. V rámci symbolismu užívaného Churchem a Tichým bychom namísto $A/A_1,...,A_n$ psali $(\alpha\alpha_1...\alpha_n)$; a namísto **V**/**T**,...,**T** bychom psali $(o\iota...\iota)$.

[3]. Podrobněji o kategoriální gramatice viz Peregrin (1998, §3.8).

[4]. Přijetí takovéto analýzy ovšem bývá mnohými autory interpretováno jako jakési odhalení toho, že v sobě analyzovaná věta ‘skrytě’ obsahuje proměnnou. Proč tohle považuji za krajně zavádějící, jsem probral na jiném místě (viz Peregrin, v tisku).

[5]. V symbolické logice existují od počátku dva poněkud rozdílné přístupy. Jeden z nich má kořeny v pracích Boola, Schrödera a Peirce; v jeho rámci je logika považována za zkoumání určitých velice obecných algebraických struktur, které lze mimo jiné aplikovat na vyplývání. Ten druhý se odvíjí zejména od Frega; a podle něj je logika *přímým* zachycováním vyplývání. Prohlásit teorie toho druhu, jaké jsme popsali zde, za *logiku* nebude činit potíže těm, kteří chápou logiku spíše v rámci té první tradice; avšak bude pociťováno jako potenciálně zavádějící těmi, kdo mají blíže k té druhé.

LITERATURA:

[1] AJDUKIEWICZ, K. (1935): Die syntaktische Konexität. **Studia Philosophica** 1, 1-27.

[2] BAR-HILLEL, Y. (1953): A Quasi-arithmetical Notation for Syntactic Description. **Language** 29, 47-58.

[3] CHURCH, A. (1940): A Formulation of the Simple Theory of Types. **Journal of Symbolic Logic** 5, 56-68.

[4] KRIPKE, S. (1963): Semantical Considerations on Modal Logic. **Acta Philosophica Fennica** 16, 83-94.

[5] LAMBEK, J. (1958): The Mathematics of Sentence Structure. **Amer. Math. Monthly** 65, 154-170.

[6] LAMBEK, J. (1961): On the Calculus of Syntactic Types. In: **Structure of Language and its Mathematical Aspects** (ed. R. Jakobson), Providence (R.I.).

[7] MOORTGAT, M. (1997): Categorial Type Logics. In: **Handbook of Logic and Language** (ed. J. van Benthem a A. ter Meulen), Elsevier / MIT Press, Oxford / Cambridge (Mass.).

[8] MORRILL, G.V. (1994): **Type Logical Grammar (Categorial Logic of Signs)**, Kluwer, Dordrecht.

[9] OEHERLE, T., E. BACH a D. WHEELER, eds. (1988): **Categorial Grammars and Natural Language Structures**. Reidel, Dordrecht.

[10] PEREGRIN, J. (1998): **Úvod do teoretické sémantiky**. Karolinum, Praha.

[10] PEREGRIN, J. (v tisku): Variables in natural language: where do they come from? In: **Variable-free semantics** (ed. M. Böttner a U. Thümmel), Secolo, Osnabrück.

[11] PRAWITZ, D. (1965): **Natural Deduction**. Almqvist & Wiksell, Stockholm.