



## Pravda a paradox: úvod do problematiky

Jiří Raclavský, Filozofická fakulta MU Brno

### Abstrakt

Tato stať je zamýšlena jako úvod do spojitě problematiky explikace pravdivosti a sémantických paradoxů, které onu explikaci pronásledují. Tyto paradoxy jsou viděny jako prostředek ukázání, že nedostatečně promyšlená explikace našeho pojmu pravdivosti (významu pravdivostního predikátu) je explikací nevyhovující. Poté jsou předvedeny rysy obou klasických přístupů, Russellovy teorie typů a Tarského hierarchie metajazyků. Nakonec jsou podány základní ideje hlavních názorových škol, které přišly poté.

### Abstract (Truth and Paradox)

The present paper is intended as an introduction to the problems connected with the explication of truth and semantic paradoxes which are affecting these explications. The semantic paradoxes are viewed as means showing that an explication of our truth-predicate which is not developed sufficiently enough is not satisfactory. The key features of both classical conceptions, Russell's theories of types and Tarski's hierarchy of metalanguages, are introduced. Finally, they are briefly characterized main contemporary approaches.

**Klíčová slova:** pravdivost; sémantické paradoxy; Russell; Tarski; Kripke

**Keywords:** truth-predicate; semantic paradoxes; Russell; Tarski; Kripke

**doi:** [10.5817/pf10-2-4](https://doi.org/10.5817/pf10-2-4)

Tato stať chce širší (spíše než specializované) filozofické veřejnosti nabídnout určitý přehled v problematice dvou úzce spjatých témat zmíněných v názvu. Exaktní vysvětlení pojmu pravdivosti je totiž pronásledováno paradoxy, což zneklidňovalo filozofy a logiky posledních více jak sta let natolik, že bylo předloženo mnoho teoretických návrhů a tak bylo prozkoumáno i mnoho aspektů problematiky. Dříve než budou charakterizovány základní problémy a trendy, bude vhodné zauvažovat nad několika obecnými otázkami.

## ***Explicace a metodologický význam paradoxů***

Není nesprávné vidět přinejmenším část práce na poli filosofie jako snahu *explikovat*. Explikovat, tedy v zájmu porozumění převést naše *pre-teoretické pojmy*, které jsou často mlhavé, neurčité, vágní, na *pojmy rigorózní*, tedy „ostré“, přesné, exaktní. To chceme proto, že s přesnými pojmy lze v usuzování pracovat s jistotou (pokud je teorie zarámována ve vágních pojmech, není jisté, co z ní vlastně všechno plyne; následně je takováto teorie sterilní, protože je lehce imunizovatelná vůči kritice). Vědci rovněž provádějí explicace, např. náš předteoretický pojem vody explikovali rigorózním teoretickým pojmem  $H_2O$ . Na vhodnou explicaci pochopitelně klademe určité metodologické požadavky. Například nemá smysl (důvod byl uveden výše) explikovat nejasný pojem pojmem právě tak nejasným. Nejen v důsledku tohoto nikdy není exaktní pojem totožný s pojmem původním. Tento rys explicace je srovnatelný s modelováním: model není shodný s modelovaným, „mapa není terén“. Pro modelování i explicaci je tedy typický prvek *idealizace*, čili některé vlastnosti pre-teoretického pojmu jsou ignorovány. Explicace i modelování mívají takovýto selektivní přístup i proto, že jsou zkoumány jen určité charakteristiky pre-teoretického pojmu. Například logická sémantika modeluje významy výrazů přirozeného jazyka nástroji logiky, kdežto lingvistika prostředky obecně jinými. Je přirozeně v zájmu věci, aby rigorózní pojem, explicans, korespondoval s oním pojmem pre-teoretickým, explicandem. Tomuto požadavku se říká *materiální adekvátnost*, věcná přiměřenost. Zde ovšem narážíme na jistý rys explicace, kterému se nelze ve všech případech explicace vyhnout. Intuice, které máme spjaty s předteoretickými pojmy, totiž ne vždy vytvářejí ve svých důsledcích konzistentní celek. Explicace, která však na budování konzistentního celku zacilována je, musí proto v některých případech z nekompatibilních intuicí volit, tedy jisté intuice nezohledňovat. Pro příklad vzpomeňme exaktní pojem ryby (tedy piscis), který nezohlednil naši intuici, že velryby jsou (velké) ryby. Jistěže je žádoucí, aby byly potlačeny pouze ty intuice, které jsou slabé, neprůkazné, minoritní. Konečně je třeba připomenout přinejmenším ještě jeden zásadní rys explicace a to její *formální korektnost*, správnost.

Tady už přicházejí na řadu paradoxy, neboť spolu s případnými kontradiktorickými důsledky jsou výmluvnými doklady toho, že nějaká explicace (či nějaká teorie na ní založená) je nesprávná. Ač se tedy na první pohled zdá, že paradoxy jsou „škodná“, mají v našem vědění svou pozitivní roli. Právě při odvolání se na pojem explicace je jasné, v čem tento *metodologický význam paradoxů* tkví. Ukazují nám totiž, že nějaká explicace je formálně nesprávná (paradoxy jsou někdy dokonce úmyslně konstruovány za tímto účelem). To nás pak přirozeně vede ke zvolení explicace jiné. Tuto jinou, alternativní explicaci je nezřídka nezbytné teprve vybudovat. Například adekvátní teorii pohybu (založenou na matematice sčítající nekonečné řady zlomků) odhalilo lidstvo o mnoho set let později, než Zenon svými paradoxy prokázal, že naše „lidová“ teorie pohybu vede k paradoxům.

Poznamenejme, že paradoxů je opravdu mnoho a že je najdeme v mnoha oblastech poznání. Obecná *klasifikace paradoxů* však ještě nebyla navržena. Relativně ujasněny jsou kategorie paradoxů *matematických* (část z nich jsou paradoxy *množinové*, byť jsou leckdy řazeny k *logickým*), *sémantických* (ty souvisí s významem a pravdivostí), *epistemologických* (ty na-

padají některé naše teorie poznání).<sup>1</sup> Ale už třeba Diderotův herecký paradox nebývá k žádné z těchto kategorií přiřazován; dále existují např. křesťanské paradoxy, které nikdo k třeba epistemologickým paradoxům neřadí; pak tu jsou i paradoxy v ekonomii, atd.<sup>2</sup>

Od antiky do dneška byly paradoxy někdy nazývány různými alternativními názvy (*insolubilia*, *sofismata*, *aporie*, *antinomie*, *kontradikce*), nicméně tyto se v současnosti neuplatňují. Připomeňme si, že název ‚paradox‘ pochází z řečtiny („para“, „doxa“, tedy „proti domnění, zdání“). Však také, jako první, antičtí Řekové paradoxy relativně bohatě diskutovali. Scholastičtí filosofové se paradoxy zabývali rovněž intenzivně (pro ně byla zajímavá sama argumentace pro to či ono řešení nějaké insolublie). Konečně třetím velkým obdobím paradoxů je právě nedávných víc jak sto let, což nepochybně úzce souvisí s nástupem teorie množin a moderní logiky.

Podle poměrně ustálené a přijímané *definice paradoxu* je paradoxem odvození, jehož – zdá se, že správně odvozený – závěr protirečí rozumně akceptovatelným předpokladům, přičemž tyto předpoklady mohou být i součástí našeho obecnějšího přesvědčení a tedy v zadání paradoxu skryté. Za *řešení paradoxu* je považováno právě dobře odůvodněné odmítnutí (diskreditace) některé z (případně skrytých) premis, které jsou nezřídka vyjádřením nějaké naší intuice anebo jsou součástí (či důsledkem) jisté teorie, vůči níž je paradox uplatňován. Pochoptitelně nám k řešení někdy stačí odhalit chybný odvozovací krok; někdy je zase diskreditován nějaký odvozovací krok.

### ***Naivní T-predikát a lhářský paradox***

Než se dostaneme k nástupu moderních paradoxů, ukážeme si, jak zásadní roli mají lhářské paradoxy (jde o celou rodinu příbuzných paradoxů) v explikaci pojmu pravdivosti. Připomeňme si, že nástup moderní logiky byl přivítán některými filozofy (jako třeba Carnapem) jako nástroj filozofické explikace a že různé otázky začaly být zkoumány na základě rozboru toho, co míníme svými vyjádřeními (epistemologický obrat k jazyku). Po období, v němž o sémantické pojmy nebyl zájem (éra nejen Carnapova syntakticismu), předložil Alfréd Tarski rigorózní pojednání (Tarski, 1956), v němž důrazně poukázal na to, že naše intuice spjaté s významem predikátu pravdivosti („být pravdivý“), *naivním T-predikátem* (málo český překlad velice často uplatňovaného anglického sousloví) vedou k paradoxu. To je závažné, neboť rigorózní explikace T-predikátu (jeho významu) by byla velice žádoucí kvůli jeho význačné roli v epistemologii. Naše intuice nás vedou k tomu, že následující schéma pokládáme za zásadní vystižení tohoto pojmu. Nechť ‚T‘ je zkratkou za T-predikát, ‚v‘ je jménem (citací) libovolné věty a ‚p‘ je ta věta sama (přeneseně je to pak propozice, význam věty), ‚↔‘ je ekvivalence (popř. pak definiční ekvivalence). Nuže tzv. *naivní T-bikondicionál* (opět málo český překlad) je:

$$T(v) \leftrightarrow p$$

<sup>1</sup> Z více knih uvádějících do problematiky paradoxů bývá ceněna kniha (Sainsbury, 1995).

<sup>2</sup> Snad bude vhodné si připomenout alespoň názvy nejdůležitějších již v antice známých sémantických (ve volnějším smyslu) paradoxů: p. hromady (vágnost), p. rohatého muže (presupozice ohledně měti rohů), Elektra (deskripce a identita), Lhář („Já lžu“, Krét’an Epimenidés: „Všichni Krét’ané jsou lháři“); právě první a poslední paradox jsou v současnosti chápány jako stále naléhavé (srov. k tomu např. Kirkham, 1995).

Slovně: věta  $v$  („Prší“) je pravdivá právě tehdy, když  $p$  (prší); neboli to, že je pravdivá věta  $v$ , obnáší, že je pravdivé to, co tato věta říká (totiž  $p$ ). Takže pravdivost vlastně tkví v *diskvotaci* (odstranění citačních uvozovek); čili *sémantický zdvih* („semantic ascent“) spočívající v citování věty a uplatnění T-predikátu je záležitostí „průhlednou“, bezproblémovou.

Jakmile ale do naivního T-bikondicionálu (či jeho slovní podoby) vložíme *lhářskou větu*  $L$ , „Věta  $L$  je nepravdivá“, dostaneme:

Věta „Věta  $L$  je nepravdivá“ je pravdivá právě tehdy, když věta  $L$  je nepravdivá. To, jak rozumíme oběma stranám této ekvivalence, nás přivádí k závěru, že její levá strana je pravdivá právě tehdy, když pravá nikoli; a naopak. Čili celá ta ekvivalence je nepravdivá, pokud je aplikována na větu  $L$ . Vzhledem k tomu, že v tomto schématu jsou volné proměnné  $v$  a  $p$  chápány jako vázané vnějšími obecnými kvantifikátory, tento celek (formule či její větná podoba) je vždy nepravdivý, neboť podmínka vyjádřená ve schématu neplatí pro všechna  $v$  a  $p$ . Takže intuitivně adekvátní explikace toho, jak chápeme T-predikát, zcela selhala.

Povšimněme si nyní určitého druhu řešení takovýchto paradoxů, jimiž jsou *ad hoc řešení*. Tato navrhuji nějaký druh omezení, pro které však nedávají přesvědčivé odůvodnění. V literatuře se kupodivu neuvádí následující přímočaré *ad hoc* řešení diskutovaného problému). Levou stranu „kondicionalizujeme“ implikací tak, že bude zamezeno destruktivní síle věty jako  $L$ ; levá strana by tedy vypadala přibližně takto:  $v \neq L \rightarrow T(v)$ . Jistě by se daly nalézt různé důvody pro odmítnutí tohoto návrhu; hlavním je však patrně ten, že takto upravený T-bikondicionál se značně míjí s naším intuitivním pojmem T-predikátu, neboť s ním si přece nespojujeme onu kondicionalizaci obsahující  $L$ .

Svědomitá řešení tedy provádějí hlubší „diagnózu“ paradoxnosti a teprve poté navrhuji „léčbu“. Diagnóza spočívá ve zjištění a rozboru – léčba pak v případném odmítnutí – premis a odvozovacích kroků paradoxu. Relativně průhledným odvozovacím pravidlům lhářského paradoxu se zde věnovat nebudeme, uvedeme si jen běžně (či méně běžně) uvažované předpoklady:

- pravdivost věty je odvislá od pravdivosti toho, co tato věta říká (od jejího významu)
- každá smysluplná věta je pravdivá či nepravdivá (*princip dvouhodnotovosti*, bivalence), ne však obojí
- citace věty je (vlastním) jménem této věty
- *autoreference* (schopnost věty vypovídat něco o sobě) je možná

Problematizovat první premisu, která latentně zahrnuje korespondenční pojetí pravdy, bývá mezi teoretiky explikace pravdivosti a paradoxů vzácné; třetí i čtvrtá premisa jsou na tom podobně. Zato druhou premisu lze považovat od 70. let 20. století za prakticky jednoznačně odmítnutou (toto samo však ještě problém neřeší).

## **Russellův paradox**

Pro začátek je vhodné podotknout, že Russell nezamýšlel předložit explikaci pravdivosti jakkoli podobnou té Tarského (jež je všeobecně velmi uznávána přinejmenším jako prvotní základ pro další, jemnější teorie). Russell na paradoxy narazil na přelomu 20. století, kdy vrcholila jeho první fáze porozumění matematice. Russellův přístup k matematice byl filosofický, nicméně z pozice amatéra se vypracoval na znalce, který dokázal kompetentně vstřebávat

matematické dění (jehož se pak dokázal i aktivně účastnit), zaobíral se zvláště teorií množin Georga Cantora. Už sám Cantor na jistý paradox narazil; zde je zjednodušená verze *Cantorova paradoxu*: mějme množinu zcela všech věcí (individuí, množin, atd.), každá množina o  $n$  prvcích má přesně  $2^n$  podmnožin (tj.  $2^n > n$ ; Cantorův teorém); množina všech věcí má  $m$  prvků, jenže přece tu pak je  $2^m$  podmnožin. Cantor takovéto „inkonzistentní multiplicity“ (např. množinu zcela všeho) potichu připouštěl (ale Russellovi věc nedala spát). Podotkněme, že Cantorova teorie množin v té době nebyla ještě jednomyslně přijímána – některé svádělo jádro Cantorovy práce (přijetí aktuálního nekonečna, ordinální čísla a jejich uspořádávání, apod.) k formulování paradoxů jako např. paradox Burali-Fortiho. Tyto paradoxy jsou zadáním složité a byly považovány buď za pomezí anomálie anebo za problémy Cantorovy teorie.

Význam *Russellova paradoxu* tak tkví kromě jiného v tom, že jeho zadání je velmi jednoduché a má navíc ještě jistou důležitou nematematickou souvislost. Zadání: uvažujme množinu všech množin, jež neobsahují samy sebe (někdy se píše: normální množiny). Obsahuje tato množina sama sebe? Známy rezultat: právě tehdy, když nikoli a nikoli tehdy, když ano. Zde je nezbytné varovat před *pseudoparadoxem holiče*,<sup>3</sup> který je Russellovi nesprávně přisuzován jako zjednodušená varianta jeho paradoxu: holič, který by měl holit sám sebe právě tehdy, když nikoli, jednoduše neexistuje (nikdo jím nemůže být). Příslušná deskripce nereferuje na žádné individuum, protože požaduje nesplnitelnou podmínku. Ta podmínka je jednoduše inkonzistentní, což má i tu souvislost, že predikát z dané deskripce sestavitelný (‘být holič ...’) má jako rozsah vždy prázdnou (nikoli jednoprvkovou) množinu. V případě Russellova paradoxu je však – při tzv. *naivní* (Cantorově) *koncepti množin* – dotyčná podmínka zcela legitimní, leč takříkajíc nepochopitelně nesplnitelná.

Roku 1902 Russell právě sepisoval knihu *The Principles of Mathematics* (Russell, 1903), když se konečně setkal s prvním dílem *Grundgesetze der Arithmetik* (Frege, 1893) matematika *Gottloba Fregeho* (hned si začali korespondovat). Russell zpozoroval, že Fregeho systém, který měl být položením pevných základů matematice (při programu tzv. logicismu), trpí na paradox, jenž dostal pak jeho jméno. Russell informoval Fregeho (Russell, 1967), který mu v uznalé odpovědi (Frege, 1967), v níž si paradox upravil do více svého jazyka, odhalil příčinu (svůj Základní zákon V.) a naplno se vyslovil o katastrofálním dosahu Russellova paradoxu pro matematiku vůbec. Fregeho chybou vlastně bylo, že předpokládal – jak bylo zcela obvyklé – že každému pojmu (tedy významu predikátu) odpovídá nějaký rozsah (byť třeba prázdný), čili že platí, že každému (smysluplnému) predikátu odpovídá rozsah.<sup>4</sup>

Frege mu se ošetření díla od paradoxu prokazatelně nepodařilo (a ani osobně se z pádu svého systému nevzpamatoval). Russell se však začal zaobírat nejen nalezením vhodného řešení, avšak i pokračováním programu logicismu (to, že matematika je odvozena z logiky),

<sup>3</sup> Tento holič holí všechny ty vesničany z vesnice (v níž žije), kteří neholí sami sebe. Existuje i varianta s *katalogem* všech katalogů, které neuvádí v sobě jako katalogový záznam samy sebe, varianta s *klubem všech sekretářek klubů*, jichž nejsou sekretářkami, varianta s *pošťákem*, který doručuje poštu všem (z vesnice, v níž žije), co sobě poštu nedoručují, ba i morbidní varianta se *(sebe)vrahem*, který zabíjí všechny ty, co nezabíjí sami sebe.

<sup>4</sup> Zde je snad vhodné připomenout jazykovou, *predikátovou variantu* Russellova paradoxu: některé predikáty jsou sobě predikovatelné (např. ‘být predikát’), některé nikoli, co ale predikát ‘být sobě nepredikovatelný’?

čehož výsledkem byla *Principia Mathematica* (Whitehead, Russell, 1910-1913). Než se pustíme do charakteristiky této vrcholné Russellovy vědecké etapy, řekněme si něco o vývoji v teorii množin. Zjistilo se – tento fakt sdělují jen zvláště znalí odborníci na paradoxy – že na Russellův paradox nezávisle narazili matematici soustředění okolo Davida Hilberta (existuje i jeden matematický paradox nesoucí jeho jméno). Hilbert svého doktoranda *Ernesta Zermela* nasměroval k teorii množin a ten roku 1908 publikoval její první paradoxů prostou axiomatizaci (ta již není naivní teorií množin). Mnohé další klasické axiomatizace obsahují analogické restriktce za účelem vyhnutí se Russellově paradoxu (jednotně se vyznačují tzv. iterativní koncepcí množin). Restriktce v těchto koncepcích jsou zhruba analogické Russellově jednoduché teorii typů (o ní dále), kterou přímo matematici nepřijali. V těchto koncepcích je vytvořena hierarchie, v jejímž nejnižším patře jsou jen množiny základních elementů, v druhém patře jsou jen množiny takovýchto množin (rovněž se připouští i přítomnost množin z prvního patra ve vyšších patrech), atp. až do nekonečna. Pro některé naivně akceptované množiny (množinu úplně všech množin apod.) tedy v takové hierarchii není místo. Patří se říci, že ne všechny matematiky tyto koncepce množin uspokojily a tak jim nezbylo než s Russellovým paradoxem zápasit nadále.

### ***Russellova teorie typů***

Russell sám postupoval poněkud jinak, jak je patrné z (narychlo sepsaného) druhého dodatku v roce 1903 vydaných *Principles* (Russell, 1903). Tato „*Doctrine of Types*“ vlastně obnáší expozici vůbec první *jednoduché teorie typů* („*Simple Theory of Types*“). Russell uvažuje, že někdy při aplikaci funkce na argument nelze dospět k přijatelnému výsledku (funkční hodnotě) a proto stanovuje obor signifikace – alternativně: *typ* – jako okruh těch objektů, pro které je daná funkce signifikantní. To vlastně znamená, že nejen funkce a objekt (ať už je tento argumentem či hodnotou té funkce) jsou separovány do odlišných typů – odlišení objektů a na nich operujících funkcí a dále druhořadových funkcí operujících na těch prvořadových funkcích, atd., jsme mohli najít už u Fregeho (Frege, 1893). Russell si ale navíc uvědomil nejen souvislost se smysluplností jako takovou, avšak navíc jeho klasifikování objektů do typů na základě signifikace odfiltrovalo i Russellův paradox. Russell si povšiml i poněkud nežádoucího rysu, jímž je ne zcela dobře odůvodněné kategorizování objektů do typů, což se přičií ideální představě, že logika by měla být univerzální (např. v oboru proměnné bychom leckdy chtěli všechny objekty, co jsou, nikoli jen objekty z určitého typu). Nejhorší ale bylo, že jistý paradox týkající se propozic nedokáže jednoduchá teorie typů vyřešit (jak Russell svědomitě informoval – Russell 1903), pokud by se nesáhlo k restrikcím ještě diskutabilnějším.

Pět let intenzivní práce v hledání optimálního řešení všech jemu známých paradoxů (při zachování možnosti vybudovat matematiku jakožto odvozenou z logiky) vedlo Russella k vývoji rámcově tří teorií, přičemž se rozhodl pro tu poslední, *rozvětvenou teorii typů* (RTT, „*Ramified Theory of Types*“). Tu publikoval poprvé ve stati *Mathematical Logic as Based on the Theory of Types* (Russell, 1908) a prakticky bez úprav byla zakomponována i do mohutných třídílných *Principia Mathematica* (Whitehead, Russell, 1910-1913) sepsaných spolu s Alfrédem N. Whiteheadem. Russell akceptoval (H. Poincarém prvotně navržený) *princip*

*bludného kruhu* (VCP, „Vicious Circle Principle“), který v jedné z variant zní: cokoli, co definuje (či nějak při tom předpokládá) nějaký „totál“, není samo prvkem tohoto „totálu“. U Russella to neznamená, že žádná množina není prvkem sebe sama: šlo mu o to, že definice nesmí být prvkem definovaného „totálu“. Russell dále technicky upřesnil, že to, co obsahuje proměnnou, nesmí být samo v oboru této proměnné. Proměnné obsahují tzv. *propoziční funkce* nebo *propozice*. Tyto propoziční funkce (potažmo propozice) nemusí být, či spíše obecně nejsou, individuovány výlučně extenzionálně, jsou obvykle tzv. intenzionální. To znamená, že Russell se vlastně odklonil od vývoje v matematice, která se striktně omezila na extenzionálně individuované funkce, funkce jakožto zobrazení (princip extenzionality: pokud funkce  $f$  a  $g$  mají pro argumenty shodné hodnoty, jsou identické).

Princip bludného kruhu vede k Russellově rozvětvené teorii typů. Ta má jakýsi nulový základ, jímž jsou individua; prvořádovými objekty jsou pak propozice a propoziční funkce, které jsou (třeba prostřednictvím proměnných) „o“ individuích; druhořádové jsou ty, které operují na těch prvořádových (typicky tedy obsahují proměnné probíhající typ prvořádových propozic či typ prvořádových propozičních funkcí), atd. Paradox lháře je pak vyřešen zjištěním, že dotyčný vyřčením ‚Já lžu‘ vlastně vyjadřuje druhořádovou propozici ‚neexistuje prvořádová propozice, kterou tvrdím a která je pravdivá‘. V souladu s principem bludného kruhu a rozvětvenou teorií typů není sama tato propozice v oboru proměnné, kterou obsahuje – takže se neděje paradoxní přisuzování pravdivosti sobě samé. Čili ‚reflexivnost‘, kruhovost (*cirkularita*) typická pro paradoxy je prý domnělá.

Reakce na rozvětvenou teorii typů byla všeobecně negativní. Pomineme-li výčitky za složitost, tak se především ukázalo, že restriktce neumožňují formulovat některé postupy v matematice nezbytné (např. tzv. Dedekindovy řezy či princip matematické indukce). Tento defekt neřešil ani Russellův *axióm reducibility*, podle kterého každé  $n$ -řádové propoziční funkci odpovídá ekvivalentní  $k$ - (např. 1-) řádová funkce. Tento (neevidentní) axióm pak jakoby činil celou rozvětvenou hierarchizaci zbytečnou a (prý) redukovatelnou na jednoduchou teorii typů. Jak první konstatoval F. P. Ramsey (Ramsey, 1925) – a po něm opakovali jiní – v matematice si zcela vystačíme se jednoduchou teorií typů (tu navrhl Ramsey, jinou L. Chwistek, nejčastěji se používá varianta A. Churche (Church, 1940); na tu Russellovu se zapomnělo). Jak po Ramseyem zdůraznil K. Grelling (Grelling, 1936), musíme odlišovat logické paradoxy (Russellův p.) a sémantické paradoxy (p. lháře), pro které je snad rozvětvená teorie typů přijatelnou metodou (toto bylo do Tarského vystoupení celkem přijímáno). Pozdější reakce na rozvětvenou teorii typů jsou ještě zápornější (je to ad hoc přístup, neřeší soudobé nekruhové paradoxy, atp.) a v současnosti je spíše předmětem přezírání.

### ***Tarského hierarchie metajazyků***

O Tarského klasické práci *Pojęcie prawdy w językach nauk dedukcyjnych* (1930, německy 1936; anglicky jako Tarski, 1956) jsme vlastně již leccos prozradili. Tarski v ní míní materiálně adekvátně a formálně správně definovat T-predikát, který by nebyl v nesouladu s korespondenční koncepcí pravdy (tato poznámka je až v Tarského filosoficky populárním textu *The Semantic Conception of Truth: and the Foundations of Semantics* (Tarski, 1944). Na počátku této rozsáhlé stati uvádí, že nehodlá podávat definici pro přirozené jazyky, neboť

tyto jsou *univerzální*, či jinak: *uzavřené*. Umožňují totiž autoreferenci a přisuzování ne/pravdivosti uvnitř sebe sama, tedy diskutování svých vlastních sémantických vlastností, což nevyhnutelně vede ke „kontradikcím“ (tehdy často užívaný název i pro paradoxy). Tarski se proto rozhodl věnovat jazykům, které mají jasně vymezenou syntax, sémantiku, a jsou též axiomatizovány, neboli pro formální jazyky (např. jazyk teorie množin); takovéto jazyky má za *otevřené*.<sup>5</sup>

Tarski při tomto odlišil *objektový jazyk*, který neobsahuje vlastní T-predikát a *esenciálně bohatší metajazyk*, který je bohatší např. o T-predikát smysluplně aplikovatelný na věty objektového jazyka. Z toho plyne, že tu vzniká *hierarchie jazyků*, z nichž každý obsahuje pouze predikáty pro jazyky nižší. Neboli T-predikát je rozdroben na (druho-jazykový) T-predikát<sub>1</sub>, (třetí-jazykový) T-predikát<sub>2</sub>, atd.

Technické detaily vlastní rekurzivní definice jsou poněkud obtížnější. Tarski definoval pravdivost na základě pojmu *splňování*, což v zásadě znamená, že otevřená formule je/není splňována určitými objekty (je/není pro ně platná), které jsou uspořádány v (nekonečné) posloupnosti, a které valuce přiřazuje volným proměnným v dané formuli. Otevřenost formulí souvisí s požadavkem rekurzivnosti definice a tudíž generováním nekonečného množství T-axiómů (srov. výše T-bikondicionál s určitou jednotlivou slovní náplní, tj. dílčí T-axióm). Důležité je, připomíná i Tarski, že pravdivost vět (formálního jazyka) je vždy relativní jen k určitému (formálnímu) jazyku. Tarski dokázal též (*Tarskéhoho*) *teorém*, podle něhož nejsou žádné sémantické pojmy definovatelné v daném jazyku (ale pouze v metajazyku k onomu jazyku).

Ačkoli Tarského konstrukce filosofy už tehdy nenadchla (což vedlo Tarského k onomu filosofickému výkladu a sebeobhajobě v Tarski, 1944), stala se okamžitě základem logického dědictví (a tím posléze byla akceptována i filosofy). A to nejen proto, že neodporovala jistému syntacticismu (jazyk je množinou formulí) – v zásadě totiž obsahuje významný pilíř všeobecně přijímané modelově-teoretické sémantiky (způsob, jak interpretovat formální jazyky, jak jim udělit množinový druh významů). Kritické reakce na Tarského formální definici však rostly zvláště v 60. let, kdy ustala spokojenost s tím, že Tarski do definice vtělil princip bivalence. Dále tomu bylo proto, že se zcela vyhnul pravdivosti v přirozeném jazyce, kde se navíc zdá být důležitější pravdivost propozic (významů vět). (Tarski se definici pravdivost propozic vyhnul, neboť tehdejší pojem propozice se mu, oprávněně, zdál dosti nejasný.) Odpor narůstal i proto, že by aplikace tarskiovského modelu na přirozený jazyk zjevně nevedly k jednomu pojmu T-pravdivosti, který v přirozeném jazyce (údajně) máme. Neméně se začala ona celá hierarchie jevit ad hoc: „hierarchizace je nezbytná, jinak vznikne paradox“ jistě není beze zbytku přesvědčujícím důvodem. Přesto všechno je Tarského hierarchický přístup respektovaným, byť „historicky pasé“, řešením.

### ***Kripke, kontextualisté, revizionisté, dialetheisté***

Do této situace, filosoficky připravené např. R. M. Martinem, přichází r. 1975 Saul Kripke se statí *Outline of a Theory of Truth* (Kripke, 1975). Definoval v ní jeden jediný T-

<sup>5</sup> Někdy se chybně tvrdí, leckdy při užití slavného Tarského úvodového příkladu (věta „Sníh je bílý“ je pravdivá =<sub>af</sub> sníh je bílý), že Tarski definoval pojem pravdy pro přirozené jazyky. To ale Tarski explicitně odmítal učinit.



predikát pro určitý formální jazyk, v němž je tento T-predikát smysluplný. Při technické konstrukci Kripke využil B. van Fraasena zpopularizovanou metodu supervaluací (a dále metodu *fixpointů*), což je proces udílení pravdivostních hodnot formulím v postupných stádiích, a silnou *trojhodnotovou logiku* S. C. Kleeneho. Tento T-predikát je tzv. *parciální*, nerozhodne totiž všechny věty jako pravdivé či nepravdivé – pro některé věty je tu totiž pravdivostní díra („gap“), podle některých je to však třetí hodnota (někdy zvaná ‚nerozhodnuto‘, ‚neuzemněno‘). Filosoficky vzato jsou vícehodnotová řešení paradoxu napadnutelná tzv. *zesíleným lhářem* („Tato věta je nepravdivá nebo má třetí pravdivostní hodnotu / má pravdivostní díru / je neuzemněná“). Také byl Kripkeho T-predikát definován v metajazyce pro daný jazyk (nikoli v objektovém jazyce samém), takže „duch Tarského“ koncepce stále zůstal. Přesto byl úspěch Kripkeho práce ohromný a bylo vytvořeno mnoho navazujících, často podobně technických prací (vzpomeňme zde aspoň S. Fefermana).

V posledních čtyřech dekáдах vytváří další školu také *kontextualisté*. Ti argumentují – ať už spíše filosoficky jako *Tyler Burge*, či spíše matematicko-logicky jako *Jon Barwise* s *Johnem Etchemendym*, či kombinovaně jako *Keith Simmons* v současnosti – že pravdivostní predikát v přirozeném jazyce není významově pevný, že jeho sémantika se mění v závislosti na kontextu. Další jsou *revizionisté*, zejména *Anil Gupta* s *Nuelem Belnapem*, kteří vybudovali různé formální systémy zachycující to, že naše sémantické pojmy jsou svébytně kruhové, přičemž při jejich uchopování uplatňujeme kruhový proces postupně revidovaných hypotéz o rozsahu těchto pojmů (obhajují kruhové definice). Nejsvráznější školou, která se rozhodla s principy klasické logiky ještě více rozejít, jsou *dialetheisté*. Zejména *Graham Priest* vehementně argumentuje pro to, že některé věty mají hned dvě pravdivostní hodnoty (tzv. „glut“), pravda i nepravda.

Kromě těchto škol nalézáme v současnosti v problematice několik solitérních osobností (např. Laurence Goldstein, který se pokouší odsoudit poslední výše uváděnou premisu lhářského paradoxu), ovšem všichni jsou zajedno v odmítnutí Russellova či Tarského přístupu. Pro současné dění je tedy příznačné to, že jsou revidovány klasické principy logiky; např. je argumentováno pro kontextovou závislost významů běžných slov, klasické definice jsou odmítnuty ve prospěch revizně-kruhových, některé věty jsou prý záraz pravdivé i nepravdivé, atd. V plejádě navržených logických systémů – obvykle určených jen k definování T-predikátu při současné absenci lhářského paradoxu – byly rovněž definovány rozmanité nestandardní implikace, nestandardní druhy negací, atd.

## Závěrem

Sémantické paradoxy jsou svou souvislostí s explikací pravdivosti a dalších sémantických pojmů natolik burcující, že snad všichni špičkoví logikové 20. století (a též mnozí analytici filosofové) neminuli tuto výzvu bez povšimnutí. Tudíž celé 20. století je bohaté na diskusi paradoxů i pravdivosti.<sup>6</sup> Byly publikovány dvě desítky specializovaných knih a přes půl tisíce studií; mnohé práce uplatňují náročné logické systémy, některé jsou spíše filosofické. Obecně má přítom úroveň bádání vysokou úroveň (výškou srovnatelnou třeba s filosofií ma-

<sup>6</sup> Připomeňme, že bylo sestaveno několik dějinám neznámých forem lhářského paradoxu a byly vyvinuty další „plnokrevné“ sémantické paradoxy (Berryho, Grellingův, aj.)

tematiky). Má se za to, že respektovatelné navrhované řešení sémantických paradoxů a teorie pravdy musí kombinovat logicko-technický i filosoficko-argumentativní prvek a platí, že řada prací toto vskutku splňuje.

### ***Vybraná literatura***

- Barwise, J. – Etchemendy, J. (1984) *The Liar*. Oxford, New York: Oxford University Press.
- Frege, G. (1893/1903) *Grundgesetze der Arithmetik*. Band I/II, Jena: Verlag Herman Pohle.
- Frege, G. (1967) Letter to Russell. In van Heijenoort, J. (ed.), *From Frege to Gödel*, Cambridge, Mass.: Harvard University Press, s. 126-128.
- Grelling, Kurt (1936) The Logical Paradoxes. *Mind*, vol. 45, no. 180, s. 481-486.
- Gupta, A.– Belnap, N. (1993) *The Revision Theory of Truth*. Cambridge, London: A Bradford Book, The MIT Press.
- Church, A. (1940) A Formulation of the Simple Theory of Types. *The Journal of Symbolic Logic*, vol. 5, no. 2, s. 56-68.
- Kirkham, R. L. (1995) *Theories of Truth: A Critical Introduction*. Cambridge, London: A Bradford Book, The MIT Press.
- Kripke, S. (1975) Outline of a Theory of Truth. *The Journal of Philosophy*, vol. 72, no. 19, s. 690-716.
- Priest, G. (2006) *In Contradiction*. 2. vyd. Oxford, New York: Oxford University Press.
- Ramsey, F. P. (1925) The Foundations of Mathematics. *Proceedings of the London Mathematical Society*, vol. 25, s. 338-384.
- Russell, B. (1903) *Principles of Mathematics*. Cambridge: Cambridge University Press.
- Russell, B. (1908) Mathematical Logic as Based on the Theory of Types. *American Journal of Mathematics*, vol. 30, no. 3, s. 222-262.
- Russell, B. (1967) Letter to Frege. In van Heijenoort, J. (ed.), *From Frege to Gödel*, Cambridge, Mass.: Harvard University Press, s. 124-125.
- Sainsbury, M. (1995) *Paradoxes*. Cambridge: Cambridge University Press.
- Simmons, K. (1993) *Universality and the Liar*. New York: Cambridge University Press.
- Tarski, A. (1956) The Concept of Truth in Formalized Languages. In: *Logic, Semantics and Metamathematics*, Oxford, New York: Oxford University Press, s. 152-278.
- Tarski, A. (1944) The Semantic Conception of Truth: and the Foundations of Semantics. *Philosophy and Phenomenological Research*, vol. 4, no. 3, s. 341-376.
- Whitehead, A. N.– Russell, B. (1910-13) *Principia Mathematica*. Cambridge: Cambridge University Press.