

K některým extravagantním teoriím významu

On Some Extravagant Theories of Meaning

Filip Tvrdý

Práce na tomto článku byla realizována na základě Grantové soutěže UP, č.p. FF_2012_024. Část textu vychází z disertační práce „Turingův test: Filozofické aspekty umělé inteligence“, kterou autor v roce 2011 obhájil na FF UP.

Abstract | Abstrakt

Semantics based on representational theories of mind has met challenges recently. Traditional accounts consider meaning as an entity with semantic properties, i.e. a mental object that denotes or represents a real-world object. The paper discusses ways of constructing meaning without representations, as shown in Rapaport's syntactic semantics and Rosenberg's eliminative theory of mind and language.

Sémantika založená na reprezentativní teorii mysli se v posledních letech setkává s problémy. Tradiční přístupy považují význam za mentální předmět, který denotuje nebo reprezentuje předmět vnějšího světa. Příspěvek se zabývá možnostmi, jak konstruovat význam bez reprezentací, a to v rámci Rapaportovy syntaktické sémantiky či Rosenbergovy eliminativní teorie mysli a jazyka.

Slovník cizích slov definuje pojem „extravagance“ jako „něco velmi zvláštního, výstřednost“ a pro jistotu uvádí i použití v poněkud pejorativním obratu „extravagantní za každou cenu“ (*Akademický slovník cizích slov* 2000, s. 216). V tomto článku se budu zabývat pohledy na sémantiku, které je možné označit za smysluplně extravagantní. Kde je prostor pro extravagantnost, musí samozřejmě existovat i zavedený standard. V sémiotice od přelomu 19. a 20. století převažuje dyadické chápání znaku, podle něhož je jeho nutnou součástí nejen označující, ale i označované, které je aspoň v případě jmen většinou tvořeno fyzickým předmětem vnějšího světa a jež Frege nazval *Bedeutung*, Peirce *object* a Carnap *extenze*.¹ V posledních přibližně patnácti letech se objevují filozofické a jazykovědné koncepce, které tento tradiční pohled na fungování jazyka zásadním způsobem revidují. Jejich kořeny spočívají v naturalistických

¹ Pro detailní terminologický přehled viz Eco 2002, s. 405.

teoriích myslí, jež vycházejí z poznatků kognitivních, výpočetních a neurálních věd a před tradičně chápanou sémantiku staví obtížně překonatelnou výzvou. Přiblížím zde přístupy dvou Američanů, počítačního lingvisty Williama J. Rapaporta a filozofa biologie Alexe Rosenberga. První by bylo možné považovat za pokus o vytvoření syntaktické sémantiky, zatímco druhý usiluje o úplnou eliminaci sémantiky jako vědecké disciplíny. Obě teorie mají za cíl vytvořit teorii jazyka, která by se obešla bez mentálních reprezentací a psychologické intencionality. Pro jejich pochopení se musíme vrátit hlouběji do minulosti, přesněji řečeno do roku 1980, kdy John Searle poprvé publikoval myšlenkový experiment čínského pokoje (Searle 1980, s. 417–424).

Searlův článek byl původně zamýšlen jako radikální odmítnutí teorie silné umělé inteligence, podle níž správně naprogramovaný počítač disponuje myslí, rozumí a má kognitivní stavy. Slavný myšlenkový experiment popisuje situaci, kdy je Searle zamčen v pokoji s velkým množstvím textu v čínštině. Neumí mluvit ani psát čínsky, dokonce ani nerozpozná čínštinu od japonštiny nebo nesmyslných klikyháků. Po chvíli dostane další text v čínštině společně se souborem pravidel pro přiřazování druhé skupiny textu k první skupině. Pravidla jsou v angličtině a Searle jim jako rodilý mluvčí rozumí. Přiřazování na základě pravidel je formální, protože symboly neidentifikuje podle významu, ale jen podle tvaru. Po další chvíli dostane třetí dávku čínského textu společně s anglickými instrukcemi, podle kterých je schopen tuto dávku textu přiřazovat k první a druhé a výsledek vrátet zadavateli. Zadavatelé pak pojmenují první dávku textu „skript“, druhou „příběh“, třetí „otázky“, symboly, které jim vrací, „odpovědi“ a pravidla v angličtině „program“. Postupem času se pravidla natolik zdokonalí a Searle získá takový cvik, že jeho odpovědi budou nerozlišitelné od odpovědi rodilého mluvčího čínštiny; budou stejně správné, jako kdyby odpovídal na anglické otázky anglicky. Přesto nerozumí čínštině a chová se stejně jako počítač: provádí počítačové operace podle formálních pravidel. Searle později učinil svůj příklad transparentnějším a srozumitelnějším, začal ještě více klást důraz na rozdíl mezi syntaxí a sémantikou:

A teď si představte, že by vás někdo zavřel do místnosti, v níž je několik košů naplněných čínskými znaky. A představte si, že (stejně jako já) nerozumíte ani slovo čínsky, že však máte k dispozici manuál pro manipulaci s čínskými znaky, napsaný ve vašem jazyce. Pravidla přitom určují způsob manipulace se znaky čistě formálně – pouze s ohledem na jejich syntax, a nikoli na jejich sémantiku. [...] Předpokládejme, že vám nějakí lidé do místnosti podají další čínské znaky a že vám poskytnou též pravidla pro vydávání znaků z místnosti ven. Předpokládejme, že znaky podávané dovnitř se bez vašeho vědomí chápou jako „otázky“ a symboly podávané ven jako „odpovědi“. Předpokládejme dále, že programátoři byli při psaní

programů tak úspěšní a váš výkon při manipulaci se symboly tak dokonalý, že vaše odpovědi brzy nebude možno odlišit od odpovědi rodilého mluvčího čínštiny. A tak jste zavřeni v místnosti, skládáte znaky a podáváte je ven jako odpověď na znaky zvenčí. Avšak na základě situace, kterou jsem právě popsal, pouze manipulací formálními znaky byste se nikdy nemohli naučit čínsky. (Searle 1994, s. 33–34)

Searle dokonce provedl formalizovaný zápis, který je depersonifikovanou a zestručnělou verzí příběhu:

Programy jsou formální (syntaktické).

Mysli mají obsahy (sémantiku).

Syntax není dostatečná pro sémantiku.

Programy nejsou myslí. (Searle 1994, s. 546–547)

Závěrem tudíž je, že ať je behaviorální projev robotů a počítačů jakkoli humanoidní, přesto nemohou jejich syntakticky vytvořené programy pochopit sémantický obsah vět, které recipují a produkují. Tato teorie byla a mnohými dodnes je považována za paradigmatickou.

Existuje mnoho způsobů, jak se Searlovým myšlenkovým experimentem polemizovat a pokoušet se o jeho falzifikaci. Zřejmě nejsilnější argument spočívá v odmítnutí poněkud vágní premisy, podle níž syntax není dostatečná pro sémantiku. Na první pohled se zdá, že tyto dvě lingvistické disciplíny spolu mají jen pramálo společného: syntax analyzuje formální vztahy slov ve větě, zatímco sémantika operuje s významem jazykových výrazů. Přesto se v nedávné době objevily názory, jež tuto dichotomii relativizují, či dokonce zcela popírají. Největším průkopníkem tohoto přístupu je zpočátku zmiňovaný William J. Rapaport (2006), který publikoval celou řadu textů popisujících vzestup sémantiky z ryze syntaktických prostředků. Rapaport považuje program za statický objekt, jenž musí být realizován aktivním procesorem. Teprve součinnost těchto dvou prvků dává vzniknout celému systému, v jehož případě už můžeme mluvit o vlastnictví kognitivních a sémantických schopností. Kromě toho ale ukazuje, že rozumění přirozenému jazyku je jen zdánlivě sémantickou záležitostí. Přesněji řečeno – rozumění je způsobeno syntaktickou sémantikou, která spočívá v počítačové manipulaci se symboly, v ničem jiném. Jako doklad pro svá překvapivá tvrzení Rapaport použil tři teze, z nichž první je nejdůležitější:

- (1) Sémantika je podle klasické definice naukou o vztazích mezi dvěma množinami entit (fyzikální objekty a jejich mentální reprezentace), zatímco syntax je disciplína o vztazích uvnitř jedné množiny (mentální reprezentace; Morris 1938). Myslicí bytosti ovšem

významy internalizují, převedou totiž vztahy mezi vnitřními pojmy a vnějšími předměty do nové skupiny interních entit. Sémantické i syntaktické vlastnosti jazykových výrazů jsou navíc realizovány týmiž konfiguracemi neuronů centrální nervové soustavy, a proto mezi nimi nemůže být principiální rozdíl.

- (2) Myslíci bytosti disponují epistemickým přístupem pouze k interním reprezentacím předmětů a nikoli k samotným objektům vnějšího světa. Můžeme být fyzikalisty v oblasti ontologie, ale musíme být senzualisty v rámci epistemologie, protože neexistuje jiný zdroj poznání o vnějším světě než empirická zkušenost.
- (3) Rozumění je rekurzivní, a proto je možné interpretovat oblast syntaxe v pojmech sémantiky. Syntaktické porozumění je elementárnější a tvoří fundament pro celý jazyk včetně sémantiky.

Rapaport má oproti Searlovi značnou výhodu, protože své názory nedokládá abstraktními myšlenkovými experimenty, ale konkrétními příklady z historie. Tvrdí, že některé lidské bytosti se ve skutečnosti nacházejí v čínských pokojích, což jim nezabraňuje naučit se používat jazyk a rozumět jeho sémantickému obsahu. Modelovým příkladem je podle něj Helen Kellerová, která ve věku několika měsíců onemocněla blíže nespecifikovanou chorobou – pravděpodobně spálou či zánětem mozkových blan – a prožila téměř celý svůj život bez sluchu a zraku. Péči své vychovatelky Anne Sullivanové se přesto vzmohla k mimořádnému intelektuálnímu výkonu a naučila se komunikovat se svým okolím pomocí prstové abecedy, což jí později umožnilo vystudovat vysokou školu a napsat několik autobiografických i jiných knih. Kellerová ve svých pamětech detailně popisuje, jak začala postupně spojovat taktilní a olfaktorické stimuly s taktilními znaky jazyka hluchoslepých. Klíčovým okamžikem bylo pochopení vztahu mezi hmatovým počítkem panenky a prstově vyhláskovaným slovem „panenka“. Ještě více přelomovým byl zřejmě moment, kdy Kellerová nahlédla souvislost mezi látkovými termíny a konkrétním prožíváním příslušných stimulací, jak se to poprvé projevilo při porozumění pojmu „voda“ (Keller 1905). Pro Rapaporta je příklad Kellerové mimořádně důležitý, protože díky svému postižení už od počátku hledala vztahy uvnitř jedné množiny taktilních zkušeností. Rapaport správně připomíná, že v jejím případě nebylo možné zřetelně odlišit lingvistickou a nelingvistickou část světa, a proto je její osvojení si jazyka ryze syntaktickým procesem. Syntax se tak může stát dostatečnou podmínkou sémantiky, třetí Searlova premisa se ukazuje být neudržitelnou.

Ještě razantnější rozchod se sémantikou ve svém díle navrhuje Alex Rosenberg, který je znám jako zřejmě nejhorlivější propagátor scientismu a redukcionismu v současné filozofii. Podle Rosenberga vůbec neexistuje *aboutness*, což je do češtiny obtížně přeložitelný termín označující intencionální vztáženost lidské mysli k nějakému předmětu nebo faktu. Rosenberg tvrdí, že celé myšlení a jazyk lze vysvětlit v termínech behaviorálních reakcí na stimuly vnějšího

světa, které jsou naprosto mechanické, a mají proto počítačovou povahu (Rosenberg 2011, s. 164–193). Jeho stanovisko má značně bizarní důsledky pro teorii významu, ale zdá se, že je podpořeno promyšlenými argumenty.

Lze si například představit jednoduchý myšlenkový experiment, který lze na rozdíl od jiných – jako je třeba ten Searlův – empiricky ověřit. V jeho rámci si představím, že se nacházím v Sydney a myslím na hlavní město Francie. Znamená to, že v mém mozku je jakási specializovaná skupina neuronů, která tuto informaci uskládá. Jak lze teoreticky uchopit vztah mezi těmito neurony a Paříží, jež je vzdálena necelých sedmnáct tisíc kilometrů od mé momentální pozice? Zdá se, že neexistuje způsob, jak takový vztah popsat fyzikálně korektně; není realizován žádnými částicemi, žádným polem, žádným vlněním, atd. Protože v něm nefigurují bosony a fermiony, což jsou podle znalostí současné vědy jediné složky, z nichž se skládá realita, pak z fyzikálního hlediska vůbec neexistuje a jiný typ existence scientistické ontologie nezná.

Jiným řešením problému by bylo tvrdit, že příslušné neurony reprezentují nebo denotují mentální reprezentace v mém mozku. Samozřejmě i ty jsou ale realizovány pomocí jiných skupin neuronů, takže se dostáváme do argumentačního kruhu. Rosenberg se navíc sugestivně ptá: jak vůbec může jeden kousíček šedé kůry mozkové *být* o druhém kousku téže matérie? Pokud chceme zůstat naturalisty, pro což máme spoustu dobrých důvodů, pak se s *aboutness* prostě musíme rozloučit. Mozek má sice ve svých neuronech zakódovány určité informace, ale ty jsou jen součástí syntaktického programu, jehož funkcí je řídit budoucí reakce organismu na podněty.

Své hypotézy Rosenberg opírá i o empirické studie, za něž byl jejich autor Eric Kandel v roce 2001 honorován Nobelovou cenou za fyziologii a medicínu. Neuropsychiatr Kandel od 60. let prováděl série pokusů na takzvaných „mořských slimácích“, tedy na plžích z rodu *Aplysia* – konkrétně se jednalo o příslušníky druhu *Aplysia californica*, což je česky zeje kalifornský. Kandel zkoumal, jakým způsobem u nich dochází k učení a ukládání vzpomínek. Zejové měří téměř půl metru, mohou vážit až několik kilogramů, ale mají vzhledem ke své velikosti velice primitivní nervovou soustavu, která se skládá jen z asi 20 tisíců mozkových buněk, oproti přibližně 100 miliardám u člověka. Na rozdíl od savců jsou jejich neurony viditelné pouhým okem, což umožňuje precizní monitorování jejich činnosti pomocí jednoduchých snímkovacích technik.

Na zejích lze provádět habituaci, senzitivizaci, ale hlavně klasické podmiňování, jež se tak osvědčilo u Pavlovových psů. Plži byli v laboratorních podmínkách učeni novým vzorcům chování, jako například reakcím na podráždění elektrickým výbojem. Kandel zároveň sledoval, jak se u zejů měnila aktivita v synapsích, a zjistil, v čem spočívá dlouho nepochopený rozdíl mezi vytvářením krátkodobých a dlouhodobých vzpomínek. Krátkodobá paměť funguje na principu

přebudování existujících synaptických spojů, zatímco při vytváření obsahů dlouhodobé paměti vznikají spoje zcela nové. Podle Kandela funguje tentýž mechanismus také u vyšších živočichů, tedy i u lidí. Veškerá mentální aktivita spojená s pamětí je tak redukovatelná na dva jednoduché mechanismy, díky nimž dochází k modifikování a vytváření synapsí. Myšlení lze explikovat pomocí ryze behaviorálních prostředků, které ovlivňují strukturu neurální soustavy živočichů. Kandel své experimenty později ověřil u potkanů a psů, ale platí rovněž u lidí – jsme stejní jako naši ostatní živočišní příbuzní, jen složitější, a máme proto větší pestrost reakcí.²

Rosenberg toto zjištění využil pro podporu tvrzení, že je iluzorní mluvit o obsahu našich myšlenek a významu jazykových výrazů. Nervové stimuly o ničem nevyprávějí, a kdo tvrdí opak, dopouští se kategoriální chyby. Rosenberg uzavírá, že mentální procesy lidí mají úplně stejnou povahu jako v případě mořských slimáků nebo počítače Watson, který byl sestaven v laboratořích IBM kvůli účasti ve vědomostních soutěžích. Introspekce a intuice nás sice přesvědčují, že jsme něco víc než plži a počítače, ale to je jen další z iluzí, jimiž nás mozek neustále zásobuje.

Pokud přijmeme Rapaportovy a Rosenbergovy závěry, nabízí se neodbytná otázka, co dále dělat se sémantikou. Domnívám se, že existuje několik východisek, jak se s dopadem počítačích a neurálních přístupů vypořádat. První možností je zachovat lingvistiku založenou na mentálních reprezentacích, ale jen na úrovni jakési *lidové sémantiky*, která by umožnila vyprávět o jazyce na úrovni zdravého rozumu. Plnila by tak obdobnou heuristickou úlohu jako lidová psychologie, jejíž pomocí predikujeme myšlenky jiných lidských bytostí, nebo naivní fyzika, díky níž dokážeme intuitivně manipulovat s okolními tělesy. Máme ovšem také druhou možnost, která spočívá v úplném opuštění reprezentacionismu jako předvědecké či nevědecké teorie. Již dnes disponují filozofové a jazykovědci nástroji, které by mohly splnit podmínky kladené na skutečně *vědeckou sémantiku*.

Sám Rosenberg například doporučuje inspirovat se intencionální sémantikou Paula Grice, kterou bychom ale paradoxně museli nejprve oprostít od intencí a přesvědčení. Zatímco Grice usiluje o převedení sémantiky na psychologii, Rosenberg ji chce ještě důsledněji redukovat až na neurologii, což je mnohem ambicióznější a hlavně dlouhodobější program. Griceovská sémantika ovšem možná není jedinou alternativou k tradiční reprezentační nauce o významu, protože se jako vhodný kandidát nabízí buď sémantika inferenční role Wilfrida Sellarse a Roberta Brandoma, nebo počítačnická lingvistika, která existuje hned v několika přijatelných variantách. Ignorovat dramatické změny, kterými kybernetika a neurofyziologie v současnosti procházejí, by ovšem pro sémantiku mohlo znamenat nebezpečí anachronického lpění na dávno vyvrácených předpokladech a následné odsunutí disciplíny na vědeckou periférii.

² Pro detailní přehled Kandelových objevů viz Kandel 2005. Populárnější shrnutí nabízí jeho nedávná autobiografie (Kandel 2008).

Literatura

- Akademický slovník cizích slov* (2000). Praha: Academia.
- ECO, U. (2002). Znaky, ryby a knoflíky. In *O zrcadlech a jiné eseje*, U. Eco, s. 392–433. Praha: Mladá fronta.
- KANDEL, E. (2005). *Psychiatry, Psychoanalysis, and the New Biology of Mind*. Washington: American Psychiatric Publishing.
- KANDEL, E. (2008). In *Search of Memory: The Emergence of a New Science of Mind*. New York: W. W. Norton.
- KELLER, H. (1905). *The Story of My Life*. New York: Doubleday.
- MORRIS, C. (1938). *Foundations of the Theory of Signs*. Chicago: University of Chicago Press.
- RAPAPORT, W. J. (2006). How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room. *Minds and Machines* 16: s. 381–436.
- ROSENBERG, A. (2011). *The Atheist's Guide to Reality: Enjoying Life Without Illusions*. New York: W. W. Norton.
- SEARLE, J. R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3: s. 417–424.
- SEARLE, J. R. (1994). *Mysl, mozek a věda*. Přel. Marek Nekula. Praha: Mladá fronta.
- SEARLE, J. R. (1994). Searle, John R. In *A Companion to the Philosophy of Mind*, ed. S. Guttenplan, s. 544–550. Oxford: Blackwell.

Klíčová slova | Key words

argument čínského pokoje, filozofie mysli, neurověda, reference, syntaktická sémantika, teorie významu

Chinese room argument, philosophy of mind, neuroscience, aboutness, syntactic semantics, theory of meaning

Filip Tvrký

Katedra filozofie FF UP v Olomouci

filip.tvrdy@gmail.com