

Posudek disertační práce Romana Chýly předkládané na FF UK, UISK, r. 2011, s názvem

Automated methods of textual content analysis and description of text structures. 165 s.

Práce napsaná anglicky (i s českými vlastními jmény – s výjimkou autorova příjmení - bez diakritik) je věnována problematice univerzálního sémantického jazyka a jeho aplikacím při automatizaci popisu struktury (nejen odborného) textu. Autor tedy řeší celý komplex úkolů rázu lingvistického, sémiotického, počítačově vědního (computer science) i informačně teoretického a praktického. Práce má osm částí, z nichž mě jako lingvistu nejvíce zajímají kap. 2, 3 a 5. První dvě se zabývají takovými základními pojmy, jako je význam, slovo a obsahová analýza textu, poslední pak hodnotí Smetáčkův analyzátor SEMAN, který autor budoval v letech osmdesátých a na jehož počátečních aplikacích jsem měl tu čest se účastnit. Čtvrtá kapitola se zabývá čtyřmi softwarovými systémy vhodnými pro aplikaci teorií sémantické analýzy textu, příp. je přímo vytvářejícími; jde o General Inquirer, Tabari, Yoshikoder a zmíněný český Seman. Z jejich srovnání pak autorovi vyplývají závěry, jak je možné ideu dekompozice významu s co největším úspěchem, a přitom účelně a elegantně realizovat při počítačové aplikaci. Úkol nesnadný, ale řekněme rovnou, že podle mého názoru se ho autor zhostil s velkým úspěchem.

R. Chýla zvolil přístup interdisciplinární, syntetizující, a tedy komplexní v míře, kterou dovoluje stav současného poznání teorií i praktických aplikací sémantické analýzy textu. Jeho znalosti literatury oboru i předmětu zkoumání jsou důkazem pečlivého výběru z množství, které může neznalého či začátečníka zavalit a zahltit. Jen kvůli zajímavosti celé problematiky bych však rád kandidátovi položil několik otázek, které by při obhajobě mohly přispět k prohloubení našeho poznání.

1. Zajímavou otázkou je možnost formalizace popisu významu a různých jeho aspektů. Počínaje Jostem Trierem a nekonče Vladimírem Smetáčkem jde při pokusech o intuitivní formalizaci a případnou komputerizaci popisu významu většinou o význam lexikální. Jde-li však o věty či výpovědi, nabízí se otázka, zda by nebylo dobré se postupně věnovat i tzv. významu gramatickému.

2. V této souvislosti se nabízí otázka po relevanci termínu *smysl*, spojeného s dalšími pojmy, jako je konotace významu nebo pragmatické funkce (nesprávně „pragmatické významy“). Mám na mysli jistý rozpor mezi aplikací metodologie obsahové analýzy, ať už v pojetí starším, novějším, nebo diferencujícím různé badatele, na jedné straně a mezi stylově neutrálním charakterem odborného textu. Jak známo, byla tato metoda analýzy kontextové sémantiky vytvořena pro sociologické, resp. politicko-vojenské účely, a šlo většinou o „čtení mezi řádky“ odhalované kvantitativními metodami zkoumání stylově a funkčně předzpracovaných publicistických textů. Z autorova předkládaného textu je zřejmý už starší poznatek, že obsahová analýza obecně má užití daleko širší, zvl. v oblasti textových komprimací, resp. i překladu. Přesto by bylo možná dobré formulovat přesně, jakou roli ve srovnání s analýzou textů publicistických tato metoda hraje primárně u textů odborných.

3. Velmi zajímavá je také otázka formalizovatelnosti některých metod původně odvozených z intelektuálního a manuálního zpracovávání textů. Tak metoda General Inquirer vzniknuvší v 60. letech 20. stol. a kol. Chýlou dobře charakterizovaná byla bez problémů komputerizována. Proti tomu např. velmi zajímavý přístup Glaser-Straussův, o kterém psal i náš Sociologický časopis v roč. 1997, se formalizaci vzpírá. Proč tomu tak je?

3.1 Kol. Chýla také poměrně kladně hodnotí přístup Anny Wierzbické v případě jejích „sémantických primitivů“, tedy našich sémů. Osobně jsem však postupu, jakými je vyvozovala, nikdy

nepřišel na chuť, srov. její formulace „dalo by se říci, že x1 by mohlo být částí x2“ nebo popis primitivu ZMĚNA pomocí formulace „vypadá jinak, než vypadala“ apod.

3.2 Kardinální otázkou lexikologie je také definice slova, které však autor věnuje jen jednu stránku. Je to přirozené, protože formalizace a komputelizace tu skýtá více možností než definice vycházející z intuitivního induktivně deduktivního myšlenkového postupu lidského. Přesto stavím tuto tezi: Pokud je lexémem míněno pojmenování jak jedno-, tak víceslovné, je nutné při strojovém zpracování předdefinovat slovníky či slovní pole. Proti tomu lingvistická teorie pracující s pojmy struktura a funkce jazykového prostředku může chápat slovo jako jednotku strukturní, kdežto pojmenování (včetně víceslovných) jako jednotku funkční.

Předkládanou práci Mgr. Romana Chýly hodnotím jako velmi zdařilou, objevnou a přesahující požadavky na disertaci k dosažení titulu Ph.D. Rád ji proto doporučuji k obhajobě.

V Olomouci 27.12.2011

Prof. PhDr. Oldřich Uličný, DrSc.

oldrich.ulicny@email.cz