

Posudok na habilitačnú prácu RNDr. Daši Munkovej, PhD.:

Evalvácia strojového prekladu

Práca Daši Munkovej je venovaná veľmi aktuálnej téme, ktorá predstavuje vo svete rýchlo sa rozvíjajúcu oblasť. Ide o jednu z mimoriadne dôležitých oblastí intelektiky, teda výskumu umelej inteligencie, ktorá sa špecializuje na vytváranie programov na spracovanie textu, vytváranie expertných systémov, na porozumenie písaného a hovoreného jazyka a jeho prekódovanie do iného jazyka.

Niet pochýb, že konkurencieschopnosť strojového prekladu voči humánnemu prekladu bude narásť a prekladateľské roboty budú postupne obsadzovať domény, kde je ich použitie v súčasnosti neefektívne alebo dokonca nepredstaviteľné. Relatívne krátke dejiny v tejto oblasti ukázali, že od počítačovej eufórie až k úplnej skepse sa vývoj uberá racionálnym smerom v tom zmysle, že očakávania a aktuálne požiadavky sú primerané súčasným možnostiam. Aj zásadný pokrok, ktorý priniesol štatistický prístup k strojovému prekladu a bol možný vďaka neporovnateľne väčšiemu objemu dát a vyššiemu výpočtovému výkonu nástrojov, ukazuje, že ďalšia cesta k zdokonaleniu strojového prekladu povedie cez spájanie hybridných metód, teda pomocou kvantitatívnych a kvalitatívnych postupov, ako je využívanie fenoménu Big Data vo forme paralelných a podobných textov, čoraz kvalitnejšie morfológicky a syntakticky anotovaných jednojazyčných a paralelných korpusov a integrovaných slovníkov a gramatík.

Problémom štatistických metód pri analýze javov prirodzeného jazyka je však to, že rôzne štatistické miery podávajú veľmi disparátne hodnoty, výsledky sa často nekryjú ani na 10%, nie je jasné, aký vzťah medzi jazykovými javmi vlastne reprezentujú a možnosti lingvistickej intervencie do procesu zisťovania dát pri predprogramovaných modeloch sú obmedzené. Výsledky štatistických metód závisia napr. aj od nastavenie prahových hodnôt signifikantnosti, teda miery filtrovania krajných, resp. periférnych hodnôt, ďalej od nastavenia veľkosti skúmaných klastrov, teda od nastavenia intervalu veľkosti skúmaného kontextu (susedný, lokálny, vetný). Problém štatistických metód je aj v tom, že štatisticky nesignifikantné môže byť jazykovo a lingvisticky relevantné.

Práca Daši Munkovej má zásadný význam, pretože ladenie nástrojov strojového prekladu na špecifiká slovenčiny za nás nespravujú iní. Tomuto výskumu by sa mala na

Slovensku venovať oveľa väčšia pozornosť a je potrebné investovať do takéhoto výskumu oveľa viac prostriedkov. Ako ukázala aj práca Daši Munkovej, mnohé problémy pri strojovom preklade sú špecifické pre konkrétnu dvojicu jazykov, preto je potrebná široká spolupráca vývojárov v oblasti strojového prekladu s čo najširším zapojením lingvistov rôznych filológií.

V súčasnosti ešte stále existujú skeptické otázky o užitočnosti vývoja komplikovaných nástrojov na strojový preklad a najmä nástrojov na jeho automatickú evalváciu, pretože vzniká nový problém hodnotenia kvality strojového prekladu.

Cenné v práci Daši Munkovej je to, že vôbec nezakrýva tento fakt a v rozsiahlej analýze histórie a vývoja v oblasti strojového prekladu, pri hodnotení jednotlivých metód sama poukazuje na nedostatky a slabé miesta. Detailne opisuje možnosti a obmedzenia rôznych metrík na hodnotenie kvality prekladu. Ukazuje sa, že štandardne používané metriky je potrebné „vyladiť“ na konkrétny cieľový jazyk. Súhlasím s autorkou, že najväčším problémom je vôbec dosiahnuť to, aby použitý analytický model meral to, čo má merať. Dôveryhodnosť nástrojov na hodnotenie kvality prekladu niekedy znižuje to, že sa samotné „mýlia“, teda označia za chybu, čo je preložené správne.

Skeptické pohľady vyplývajú niekedy aj z faktu, že zložitý a výpočtovo náročný model hodnotenia strojového prekladu odhalí chyby, ktoré sú bez problémov viditeľné na prvý pohľad voľným okom a vzniká tak dojem, že ideme – ľudovo povedané – s guľometom na vrabce.

Hlavným problémom matematických analytických modelov evalvácie je však to, že dosadzované premenné, reprezentované konkrétnymi hodnotami sa opierajú o vágne pojmy, ktorými sú označované stupne škály napr. pri definovaní adekvátnosti pomocou pojmov celá/väčšina/veľa/málo/žiadna alebo plynulosti pomocou pojmov bezchybná/dobrá/neprirodzená/neplynulá/nezrozumiteľná. Problémom môže byť aj korelácia medzi týmito dvomi kritériami. Napríklad plynulosť ešte nemusí automaticky korelovať s adekvátnosťou. Pri manuálnom hodnotení kvality hodnotenia vstupuje do hry aj subjektívny faktor, a to samotná kvalita a momentálna dispozícia hodnotiteľa.

Súčasná hlavná metóda strojového prekladu prostredníctvom identifikácie fráz prekonala problém izolovaného prekladu jednotlivých slov a rešpektuje jazykový fakt, že každé slovo má svoj kolokačný potenciál, ktorý nikdy nie je nekonečný. Preto sa napr. aj korpusová lingvistika a v rámci nej kolokačná gramatika sústreďuje najmä na odhalenie prirodzenej kolokability prvkov, ktorá je často arbitrárna, jazyková a nemožno ju vysvetliť na

základe sémantiky slov. Možno tak čiastočne odstrániť problém ambiguity jazykových prvkov a polysémie a presnejšie identifikovať ustálenú kombinatoriku. Problémom však je, že sa ako významné často identifikujú nesúvisiace nesyntagmatické konkurencie. Tento problém je výrazne prítomný aj v strojovom preklade, preto považujem za dôležitý oveľa hlbší výskum v oblasti kolokácií, ktorý by umožnil integrovať do prekladateľských nástrojov aj tzv. kolokačné slovníky. Situáciu komplikuje však fakt, že ustálené vzťahy medzi jazykovými jednotkami sú v texte často narušované aktualizacími prvkami, časté sú transformácie, porušovanie tzv. neutrálneho slovosledu atď. Vývoj bude preto smerovať k vývoju nástrojov, ktoré dokážu identifikovať vzťahy na úrovni vety, čo naznačuje aj autorka. Tam prirodzene prídu problémy s identifikáciou vzťahov v rámci nadvetnej syntaxe, ako je napr. rozpoznanie deiktických, anaforických a kataforických vzťahov atď.

Analýzu strojového prekladu a aplikáciu nástroja na jeho automatickú evaluáciu v práci Daši Munkovej možno považovať za sondu, ktorá poskytuje základ na formulovanie ďalšieho výskumu v tejto oblasti. Autorka sa venuje najmä otázke súladu automatických metrick strojového prekladu s ručným hodnotením. Dôležité však bude overiť efektívnosť metódy na oveľa väčšom materiáli a samozrejme aj na rôznorodejších textoch. Aj keď bol materiálom východiskom len malý text s 37 vetami, poskytoval už istý priestor na vyhodnotenie kvality samotného evaluačného nástroja. Tu by sa žiadala hĺbková lingvistická analýza správnosti automatického hodnotenia. Autorka sa pri zdôvodňovaní príčin kvality a chýb hodnotenia strojového, posteditovaného a humánneho prekladu odvoláva na tzv. variabilitu jazyka. Je to veľmi vágne zdôvodnenie príčin chýb v preklade alebo pri jeho hodnotení. Tento pojem, ktorý používa autorka v práci, by si práve zaslúžil exaktnejšiu analýzu typov chýb v preklade a aj v samotnom hodnotení prekladu. Na základe toho by sa dali aj exaktnejšie definovať typy chyby a posuny tak v preklade, ako aj pri jeho automatickom hodnotení. Keďže pri manuálnej evaluácii kvality prekladu sa môžu používať rôzne techniky (napr. blind post-editing, explicit marking of errors, Quiz-Based Evaluation a i.), bolo by vhodné aj porovnanie typov chýb hodnotiteľov vzhľadom na použitú metódu. Je to dôležité aj preto, že pri manuálnom editovaní prekladu, ale aj pri manuálnom hodnotení typov chýb nie je celkom možné reprodukovать postupy a techniky evalvátorov, medzi ktorými býva často aj nízka zhoda v hodnotení. V práci mi chýba práve analýza typológie chýb. Uvedomujem si však, že krátky text nedával priestor na takéto zovšeobecnenia.

Celkovo hodnotím prácu Daši Munkovaj ako novátorskú a v našich pomeroch priekopnícku. Prináša bohaté informácie s početnými referenciami a odkazmi na vývoj a súčasný stav v oblasti strojového prekladu a jeho automatickú evalváciu, objektívne hodnotí súčasný stav v tejto vednej oblasti, aplikuje najnovšie poznatky a metódy s použitím vlastného analytického nástroja. Autorke odporúčam, aby z tejto problematiky napísala aj vysokoškolskú učebnicu pre študentov translatológie.

Odporúčam, aby bol RNDr. Daši Munkovaj, PhD. po úspešnej obhajobe udelený titul docent v odbore 2.1.35 Prekladateľstvo a tlmočníctvo.



V Bratislave, 7. 4. 2014

Prof. Peter Ďurčo, CSc.