

Oponentský posudek diplomové práce

Bc. Klára Kopecká

Kategorizace úprav strojového překladu při posteditaci: jazyková kombinace angličtina – čeština

Praha 2021, 84 s. + přílohy.

vedoucí práce: **PhDr. Bc. Tomáš Svoboda, Ph.D.**

oponent práce: **Mgr. Věra Kloudová, Ph.D.**

Předkládaná diplomová práce Kláry Kopecké je teoreticko-empirickou studií zabývající se posteditací strojového překladu (MTPE) ve směru z angličtiny do češtiny. Předmětem zkoumání jsou úpravy provedené posteditory a změny provedené během následné revize. Autorka diplomové práce klasifikuje a kategorizuje zjištěné změny a pokouší se zmínit zásadní aspekty provázané s touto oblastí (včetně zadání, jež posteditoři dostali, různých úrovní posteditace, zkušenosti posteditorů). Kromě samotného výzkumu je cílem práce zamyslet se nad možným kurzem věnovaným MTPE v rámci vysokoškolského studia.

Pozitivně hodnotím provázanost s překladatelským trhem (spolupráce se společností Memsource, kde je autorka diplomové práce zaměstnána, výzkum se týká datasetu poskytnutého touto společností) a aktuálnost výzkumu (na trhu byl zaznamenán prudký nárůst zájmu o texty vzniklé posteditací strojového překladu).

Teoretická část

V úvodní části se jedná o promyšlený a dobře strukturovaný přehled historie různých typů strojového překladu, avšak u toho nejdůležitějšího (NMT) chybí (kromě jedné nijak blíže neokomentované citace z diplomové práce) podstatnější vysvětlení jeho fungování a odkazy na relevantní literaturu (např. Philipp Koehn, *Neural Machine Translation*, Cambridge University Press 2020), namísto toho zde najdeme odkaz na popularizační článek ze serveru iDnes z roku 2017 (Kasík). Další popularizační články od téhož autora i od autorů dalších (Novák) jsou citovány i dále u přehledu jednotlivých překladačů. Z vědeckého hlediska je třeba pečlivě zvážit, zda jsou uvedené informace spolehlivé. Pokud už se v diplomové práci má něco takového objevit, je záhodno zvolit např. poznámku pod čarou. Pro obhajobu doporučuji se v tématu zorientovat a podat přehled o současné situaci v této oblasti.

Zdařile shrnuty jsou současné trendy v oblasti strojového překladu, někde však chybí ozdrojování (TER, BLEU) (k BLEU se nabízí např. Markus Freitag, David Grangier, and Isaac Caswell. 2020. BLEU might be guilty but references are not innocent. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 61–71, Online. Association for Computational Linguistics). Autorka ovšem na str. 19 správně zmiňuje nevýhody této metriky. Uvítala bych případně také zdroj ke zmiňované metrice chrF (Maja Popović 2015: <https://aclanthology.org/W15-3049/>).

Pozitivně hodnotím, že si autorka mj. všímá problematiky hodnocení překladu, z novější literatury doporučuji pro obhajobu nahlédnout do následující studie: <https://arxiv.org/pdf/2104.14478.pdf> (Experts, Errors, and Context: A Large-Scale Study of Human Evaluation for Machine Translation, Markus Freitag et al., 2021). Náležitá je také zmínka o document-level MT, který bude v dalším bádání v této oblasti velice důležitý (doporučuji: Lopes et al. 2020: Document-level Neural MT: A Systematic Comparison).

Další úseky teoretické části jsou relevantním shrnutím problematiky, o tématu posteditace je pojednáno (pro účely této DP) dostatečně. Velmi cenný je přehled jednotlivých pracovišť, kde se v překladatelských kurzech pojednává rovněž o problematice nástrojů CAT a strojového překladu, resp. posteditace.

Praktická část

Praktická část obsahuje analýzu datasetu (jenž nevznikl primárně pro účely této práce, ale byl řešitelce poskytnut firmou Memsources) a návrh kategorizace posteditačních změn z hlediska lingvistiky. (Směr angličtina – čeština, texty přeloženy 13 strojovými překladači (Google Translate, Microsoft, 11 verzí překladače CUBBITT)). Posteditaci provádělo 15 posteditorů za předem daných podmínek (včetně měření času stráveného nad zadaným úkolem). Následnou revizi provádělo 17 jiných pracovníků. Obě skupiny vyplnily dotazník týkající se jejich zkušenosti s překlady a posteditací. Výstupy včetně dotazníku byly autorce této práce předány, nezajišťovala je tedy sama, pro svou analýzu si vybrala pouze výsek datasetu (od 8 posteditorů, každý pracoval s 8 stejnými zdrojovými texty, avšak jinými výstupy v cílovém jazyce).

Vlastní analýza: Autorka správně na str. 77 poznamenává, že *způsob kategorizace, definice jednotlivých kategorií i zařazování jednotlivých úprav do lingvistických kategorií je do jisté míry subjektivní* a navrhuje případný budoucí výzkum spočívající v tom, že posteditoři uvedou také motivaci pro dané změny. S tím nelze než souhlasit. (Některé kategorie, resp. jejich zařazení jsou opravdu k diskusi, např. interpunkce, pravopis, slovosled a gramatika jsou společně v rámci kategorie plynulosti – a podobných případů zde nalezneme více. Slovosled tedy „není gramatika“? Nepovažuji to za vyložení chybné, slovosled opravdu „není vždy gramatika“, spíše je to poněkud nápadné, pokud zde není uvedena motivace pro danou volbu a vysvětlení, jak autorka při kategorizaci postupovala. Zřejmě se inspirovala hned několika pohledy, a to jak výzkumem strojového překladu, tak výzkumem lingvistickým a translologickým – vznikl však poněkud nepřehledný a nepřehledný systém.)

Co v této části práce zásadně postrádám, je alespoň nějaká odborná opora pro návrh dané kategorizace, odkazy na relevantní literaturu, kde se autorka případně inspirovala. *Na počátku analýzy jsme si definovali základní skupiny úprav, ke kterým jsme očekávali, že bude docházet, např. úpravy slovosledu, syntaxe nebo terminologie* (str. 51): ale podle čeho? V tabulkách chybí (podle mého názoru oprávněně očekávaný) komentář k jednotlivým změnám.

Teoretická část je tedy poměrně zdařilá (až na výše uvedené výtky), praktická část poněkud postrádá teoretické ukotvení (minimálně diskusi o možných přístupech a proč byl nebo nebyl zvolen některý z nich) a komentář z pohledu autorky. Zdá se, že autorku tlačil čas (?), a proto z velmi cenných dat nevytěžila tolik, kolik by se dalo očekávat.

Menší výhrady mám rovněž k formálním chybám a nedostatkům: chybná interpunkce (chybí čárka za *Ph.D.*, str. 3), četné formulační nedostatky v abstraktu (např. *Práce překladatele dnes neznamená jen převádění obsahu z jednoho jazyka do druhého za pomoci vlastních znalostí dvou jazyků a obsahu*, str. 4; *porozumění zdrojového textu*, str. 7; *systému založeném*, str. 15; chybějící interpunkce na str. 15 i jinde; *škálu shod strojové překladu*, str. 18; spojovník místo pomlčky, str. 28). Také styl je na mnoha místech neadekvátní (*Na počátku analýzy jsme si definovali základní skupiny úprav, ke kterým jsme očekávali, že bude docházet*, str. 51), a to včetně toporné syntaxe připomínající skladbu anglicky psaných vět nebo redundantních výpovědí jako *Nyní popíšeme jednotlivé systémy* (str. 13).

Na základě výše uvedeného doporučuji diplomovou práci Kláry Kopecké k obhajobě a navrhuji ji ohodnotit známkou **velmi dobře**, o konečném výsledku rozhodne obhajoba.