

Oponentský posudek diplomové práce

Kamil Kos: Rich Features in Phrase-Based Machine Translation

Popis práce

Předložená práce se po vyložení základních principů statistického strojového překladu zabývá dvěma přístupy ke zlepšení kvality jeho výstupu. První přístup popsáný ve třetí kapitole spočívá v použití alternativní metriky SemPOS při optimalizaci parametrů překladového systému Moses (namísto obvyklé metriky BLEU). Druhý přístup zdokumentovaný v kapitole 4 je zaměřen na obohacení log-lineárního modelu o řadu rysů využívajících znalost kontextu překládané jednotky na straně vstupního jazyka. Ani jedna z cest nevedla k dosažení statisticky signifikantního zlepšení, možné příčiny jsou diskutovány v kapitole 5. Za shrnující šestou kapitolou následuje seznam použité literatury, příloha s výčtem použitých sémantických slovních druhů a příloha s uživatelskou dokumentací nového rozhraní k nástroji Z-MERT pro hledání parametrů log-lineárních modelů.

Včetně seznamu literatury a příloh má práce 72 stran. Práce je psána anglicky. Součástí práce je přiložený CD-ROM se zdrojovými kódy a daty použitými v experimentech.

Hodnocení

Práce je přehledně strukturovaná. Po jazykové stránce je text velice kvalitní, množství chyb, které jsem našel, je nepatrné. Autor prokázal, že dobře rozumí moderním metodám statistického strojového překladu, má široký přehled o aktuální literatuře, je schopen v tomto oboru navrhnout a realizovat řadu dobře motivovaných experimentů a řešit související implementační obtíže.

První série experimentů vychází z předpokladu, že hledání parametrů systému Moses by mohlo být zlepšeno použitím autorovy metriky SemPOS, která při srovnávání kvality výstupu různých překladových systémů koreluje s lidských hodnocením o něco lépe než standardní metrika BLEU. Tento předpoklad bohužel experimenty nepotvrdily. Na první pohled je překvapivé hlavně to, že při optimalizaci na metriku SemPOS se úspěšnost výsledného systému zhoršila jak v metrice BLEU, tak v samotné metrice SemPOS. Autor to vysvětluje tím, že SemPOS zanedbává povrchové formy a z využívá pouze lemmata. Dalším, podle mého názoru významnějším důvodem může být ale nezapočítávání funkčních slov a z toho plynoucí narušení spolupráce překladového a jazykového modelu během dekódování. Stabilita trénování pomocí SemPOS může být narušena rovněž tím, že nevážené průměrování přes jednotlivé sémantické slovní druhy (vzorec 3.2) fakticky vede ke značnému zvýšení významu méně častých sémantických slovních druhů.

Třetí kapitola nese název „Rich Annotation in MT Evaluation“, který je ale podle mého soudu poněkud zavádějící. Ve skutečnosti se z veškeré představitelné a realizovatelné anotace používá pouze lemmatizace, rozlišení slov na funkční a plnovýznamová a kategorizace autosémantických slov do jednotlivých sémantických slovních druhů. Navíc přínos hodnot sémantických slovních druhů zde bude podle mého názoru minimální, protože jsou ve většině případů determinovány lemmatem (nebo bude dokonce sporný, viz výše). Ať už tomu budeme říkat bohatá anotace, nebo ne, bylo by jistě zajímavé zjistit, zda by vyvinutou metriku nebylo možné modifikovat tak, aby pracovala pouze s lemmaty.

Druhé téma zkoumané v přeložené práci vychází z rozumného předpokladu, že kvalitu překladu by mohlo zvýšit zapojení informace o kontextu překládané jednotky do log-lineárního modelu. Sice to narušuje původní koncepci SMT jakožto noisy-channel modelu (ten pracuje pouze se zpětným překladovým modelem, pro který je zbytek vstupní věty neznámý), nicméně úspěšné experimenty zapojující dopředný diskriminativní překladový model již byly publikovány. Bohužel ani zde se nepodařilo přes řadu přidanych rysů dosáhnout signifikantního zlepšení. Autor se zde také snaží podat vysvětlení, možnou příčinou může být opravdu nevhodnost optimalizační metody MERT pro hledání v prostoru s větším množstvím parametrů.

Autor zmiňuje, že nejlepšího zlepšení bylo dosaženo po přidání rysů získaných závislostní analýzou věty. Dovolil bych si požádat o vysvětlení motivace pro třetí rys (relativní pozice uzlu vůči kořeni věty). Vzhledem k rekurzivní povaze syntaktických struktur mne nenapadá žádná jazykově relevantní interpretace nebo možný přínos takové vlastnosti.

V práci vysoce hodnotím zejména pečlivé zdokumentování a vyhodnocení experimentů a snahu o rozbor příčin neúspěchu.

Závěr

Autor prokázal, že dokáže samostatně řešit náročné výzkumné téma. Předložená práce je zpracována velmi pečlivě a s rozhledem v oboru statistického strojového překladu. Doporučuji ji k obhájení.

V Praze, 27. srpna 2010

Zdeněk Žabokrtský
Ústav formální a aplikované lingvistiky
MFF UK, Praha