

Název práce: Bohaté rysy ve frázovém strojovém překladu

Autor: Kamil Kos

Katedra (ústav): Ústav formální a aplikované lingvistiky

Vedoucí diplomové práce: RNDr. Ondřej Bojar, Ph.D.

e-mail vedoucího: bojar@ufal.mff.cuni.cz

Klíčová slova: strojový překlad, hodnocení kvality, kontextový model, suffixové pole

Abstrakt:

V této práci zkoumáme metody, jak zlepšit kvalitu statistického strojového překladu použitím bohaté lingvistické informace. Nejdříve popíšeme SemPOS – metriku, která využívá mělké sémantické reprezentace vět k hodnocení kvality strojového překladu. Ukážeme, že i když tato metrika dosahuje vysoké korelace s lidskými hodnoceními kvality překladu, není samostatně vhodná pro optimalizaci parametrů systémů strojového překladu. Za druhé rozšíříme základní log-lineární model používaný ve statistickém strojovém překladu o kontextový model zdrojové věty, který pomáhá lépe rozlišovat mezi různými možnostmi překladu dané fráze a pomáhá vybrat nejvhodnější překlad pro daný kontext v aktuální větě.