

Posudek oponenta bakalářské práce

Student: Lipták Šimon

Téma: Slovenská lemmatizace (id 18337)

Oponent: Dytrych Jaroslav, Ing., UPGM FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
2. **Splnění požadavků zadání** zadání splněno pouze částečně

Vytvořený systém nenabízí možnost generovat správné tvary slov pro daný větný kontext. Možnosti automatického získávání morfologické informace z korpusových textů nejsou diskutovány.
3. **Rozsah technické zprávy** je v obvyklém rozmezí

Práce má 32 normostran textu a cca 4 normostrany obrázků.
4. **Prezentační úroveň předložené práce** 70 b. (C)

Práce má logickou strukturu a jednotlivé kapitoly na sebe navazují. Některé části teorie jsou dle mého názoru rozepsány až příliš do detailů, zatímco samotnému technickému řešení je věnován malý prostor. Preferoval bych, kdyby cíle práce byly stanoveny ještě před teoretickým rozbohem.
5. **Formální úprava technické zprávy** 55 b. (E)

Práce je psána slovensky, takže nemohu detailně posoudit jazykovou stránku. Nicméně je zde větší množství zjevných překlepů a chyb interpunkce. Z typografického hlediska jsou zde rovněž chyby, např. jednoslabičné a neslabičné spojky a předložky na koncích řádků a horší čitelnost popisků v grafu 4.2.3. Čtvrtou úroveň číselovaných nadpisů nepovažuji za vhodnou. Tabulky a obrázky nejsou odkazovány z textu, ale jsou v něm pouze umístěny pod částmi, ve kterých je autor zmiňuje.
6. **Práce s literaturou** 75 b. (C)

Výběr papírové literatury i elektronických zdrojů hodnotím jako nadprůměrný. Byly zvoleny vhodné zdroje, které dobře pokrývají dané téma. Negativně však hodnotím fakt, že obrázek 2.3 je převzatý bez řádné citace (pravděpodobně z Wiki Českého národního korpusu) a citování některých dalších zdrojů je na hranici korektnosti.
7. **Realizační výstup** 40 b. (F)

Vytvořený lemmatizátor i stemmer fungují. Lemmatizátor je pomalý, což student v práci uvádí a zdůvodňuje. Obdobně jsou v práci uvedena i další omezení a navíc zde není možnost generovat správné tvary slov pro daný větný kontext, což bylo požadováno zadáním. Vytvořené nástroje fungují pouze v interaktivním režimu, což značně snižuje jejich další využitelnost.
8. **Využitelnost výsledků**

Dle mého názoru vytvořené řešení není bez dalších úprav využitelné.
9. **Otázky k obhajobě**
 - Řešil jste možnost generovat správné tvary slov pro daný větný kontext?
 - Proč jste neprovedl žádnou optimalizaci práce se slovníkem pro zvýšení rychlosti?
 - Proč lze vytvořené nástroje využívat pouze v interaktivním režimu?
 - K čemu lze Vámi vytvořené nástroje prakticky využít?
10. **Souhrnné hodnocení** 50 b. dostatečně (E)

Technická zpráva obsahuje nezanedbatelné množství formálních nedostatků. Student se zaměřil především na teoretickou stránku práce, ale realizované řešení není optimální, má řadu omezení a možnost jeho dalšího využití je diskutabilní. Zadání nebylo zcela splněno (2. a 5. bod byly splněny částečně).

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 1. června 2016

.....
podpis