

Posudek bakalářské práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

posudek vedoucího

Autor/ka: Matěj Korvas

Název práce: Empirické meze vybraných modelů strojového překladu

Studijní program a obor: Informatika, Obecná informatika

Rok odevzdání: 2010

Jméno a tituly vedoucího/ponenta: RNDr. Ondřej Bojar, Ph.D.

Pracoviště: ÚFAL MFF UK

	e x c e l e n t n í	o d p o v í d a j í c í	s l a b š í	n e v y h o v u j í c í
Náročnost zadaného tématu	X			
Míra splnění zadání		X		
Rozsah práce		X	X	
Struktura textové části práce	X	X		
Analýza				
Vývojová dokumentace				
Uživatelská dokumentace		X	X	
Jazyková a typografická úroveň				
Návrh a design implementace				
Kvalita zpracování softwarové části				
Stabilita aplikace				

Nejvýznamnější klady:

Matěj Korvas se dokázal vypořádat s velmi široce formulovaným zadáním. Samostatně našel a navrhl míry vhodné pro empirické studium modelů strojového překladu.

Práce je dobrým počátečním bodem pro další výzkum, provedená měření otevírají závažné otázky, např.:

- Proč překlad s morfologickou analýzou (označován jako Afact) při snížení OOV vede k nižšímu počtu dosažitelných vět? (Je to způsobeno překročením limitu možností překladu, ttl?)
- Jak lépe využít hloubkovou informaci (Tfact)? Z prvních výsledků se zdá, že ztěžuje úlohu ve více ohledech: OOV, entropie rozhodnutí. Jednou z kompenzací je dle mého nižší počet jednotek, což však Matějovy míry nepostihují.

Nejzávažnější nedostatky:

V návrhu míry entropie rozhodování (2.3) jednak chybí motivace (v praxi se užívá beam-search, proto vyšší entropie vede k vyššímu riziku chyby v prohledávání), a zejména vůbec není postížena entropie volby segmentace vstupu na překládané jednotky a entropie volby pořadí přeložených jednotek. Na toto zanedbání bylo třeba přinejmenším upozornit.

Pro závěrečný spěch jsou struktura i formulace na některých místech málo uhlazeny, namátkou: nezavedené termíny v tabulce 1.1 bez dopředného odkazu (tabulka sama je velmi užitečná); oddíl 2.1 odstavec 1, kde je patrné, že text byl původně uspořádán jinak; velikost testovací množiny vět není uvedena v kap. 3, ale (bez včasného odkazu) až v příloze.

Další poznámky:

Extremní model překladu po písmenech lze při použití frázového překladu velmi dobře motivovat: systém se sám může naučit, která slova překládat v jednom kroku (“fráze písmen”), a která skládat ze součástek jako kmen+koncovka. V Matějově implementaci model pravděpodobně selhává kvůli chybám v zarovnání tokenů — to mělo být odhadováno tak, aby pokud možno respektovalo zarovnání slov.

Proč jsou v IV-n v položce “Times Used” zahrnuta i trénovací data?

	v ý b o r n ě	v e l m i d o b ř e	d o b ř e	n e p r o s p ě l / a
Návrh známky		X		

Datum: 27.8.2010

Podpis:

