

Posudek diplomové práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

Autor/ka: Odřej Kazík

Název práce: Lingvistická komprese textu

Rok odevzdání: 2009

Cílem práce byl návrh a implementace speciálního kompresního algoritmu určeného pro kompresi textů v přirozeném jazyce. Součástí zadání je i poměrně konkrétní popis toho algoritmu. Jedná se o práci na pomezí několika oborů.

Posouzení výsledné práce považuji za obtížné z několika důvodů. Na jedné straně je poměrně zřejmé vidět značné úsilí, které řešitel v souvislosti s touto prací vynaložil. Na straně druhé toto úsilí není dostatečně reflektováno v textu práce. V průběhu řešení došlo k mnoha důležitým a na základě osobní konzultace s řešitelem mohu říct, že i kvalifikovaným rozhodnutím. Tato rozhodnutí však nejsou v textu dostatečně dokumentována, důležité úvahy a analýzy, které proběhly, v textu chybějí (nutnost implementace vlastního taggeru, volba značkovací metody pro NNTagger, vliv variace větných typů na kvalitu komprese na str. 51 a jiné). Řešitel v textu na několika místech vysvětluje možné alternativy, ze kterých volil, odůvodnění jeho volby však chybí. S tím souvisí i z mého pohledu nedostatečné vyhodnocení výsledků této práce. Výsledkem práce je velmi sofistikovaná a komplikovaná metoda, z textu práce však není poznat, nakolik je komplikovanost metody odůvodněná.

Dále mám jisté výhrady k formě textu práce. Text trpí jistou nekonceptností, mnohé kapitoly na sebe příliš nenavazují a čtenář se občas ztrácí v textu, který není navázaný na téma práce. Některé kapitoly by bylo vhodnější umístit mezi dodatky. Čitelnosti by velmi prospělo zvýraznění formálních pojmů v textu. Na některých místech postrádám referenci k odkazovanému textu (Str. 39 nahoře: přidat odkaz na sekci 6.1, Str. 54: *model slov* – viz 6.2.3).

Největší plus práce vidím v její mezioborovosti. Řešitel prokázal hluboké znalosti nejrůznějších témat. Ve všech bodech splnil zadání práce a přidal mnoho i nad jeho rámec. Za prakticky provedenou práci však poněkud zaostává odevzdaný text.

Práci doporučuji k obhajobě, hodnotím ji pozitivně a domnívám se, že by na jejím základě mohla vzniknout zajímavá vědecká publikace.

Datum: 28.8.2009

Jan Dědek

Podpis:

