

Diplomová práce – posudek vedoucího

David Bertoli: Kompresní korpus

Diplomová práce splnila svůj cíl, vytvoření nástroje pro snadné testování kompresních programů a jejich různých nastavení nad vybranými kolekcemi dat. Změřené výsledky se uchovávají pro další zpracování, je umožněno jejich porovnání s jinými programy či jinými nastaveními programů a také statistická analýza získaných dat pomocí několika typů testů.

Kompresní korpus je řešen formou dvou webových aplikací (jedné pro administrátora, druhé pro uživatele) a datového úložiště, které je schopné fungovat jak na operačních systémech typu windows, tak i unix. Oceňují, že instalace probíhá pomocí spuštění jediného souboru, který si sám dohledá a s minimálními zásahy uživatele nainstaluje potřebný software (java, tomcat, postgres)

První část práce se zabývá návrhem struktury souborů v kompresním korpusu. Cílem je zahrnout všechny běžně zkoumané kategorie dat. Struktura korpusu je hierarchická, což zvyšuje variabilitu možností, nad kterými typy dat provádět testy.

Administrátorská aplikace umožňuje například: měnit hierarchii korpusu, přidávat soubory, přidávat a editovat uživatele.

Pomocí uživatelské aplikace lze nahrávat na server kompresní programy pro testování (včetně programů, které nejsou tvořeny jen jedním souborem), určovat masky parametrů, se kterými se program spouští, vybírat na jaké části korpusu mají testy proběhnout a automatické pouštění testů. Z mnoha dalších funkcí by jsme zmínil export výsledků do tabulek v TeXu či zdrojových kódů pro vykreslení grafů v programu R.

Myslím si, že vytvořená aplikace je použitelná pro účel, pro který byla zamýšlena. Pro začínajícího uživatele je velmi těžká na ovládání, nicméně uživatelská dokumentace je kvalitní a lze z ní ovládání pochopit.

Předloženou práci považuji po všech stránkách za práci splňující kritéria pro diplomové práce na MFF UK a doporučuji ji k obhajobě.

Praha, 20. 12. 2008

Mgr. Jan Lánský, KSI MFF UK

