

Oponentský posudek na diplomovou práci

Evaluation of Mathematics Retrieval

Autor práce: Martin Líška, student Fakulty informatiky MU

Student se ve své práci zabývá problematikou vyhledávání v matematických formulích v odborných matematických textech. Navazuje zároveň na svoji vlastní bakalářskou práci, která se věnuje stejnému tématu.

Práce začíná úvodem a druhou kapitolou – v její první části přehledem historie a evaluačních metod v oblasti information retrieval (IR), ve druhé části následuje podrobný rozbor problémů a řešení IR a evaluace v oblasti matematiky (mathematical IR, dále MIR). Třetí kapitola shrnuje současné softwarové systémy pro MIR, jejich klady a zápory. Ve čtvrté kapitole je podrobný rozbor vytvořeného systému MIaS, postupů a algoritmů, které používá (indexace, vyhledávání, webové rozhraní). Poslední kapitola je věnována evaluaci vyhledávání s pomocí systému MIaS. V přílohách je stručný přehled přiložených souborů (software MIaS), diagramy zobrazující architekturu systému MIaS (komponenty, třídy) a kopie evaluačních dokumentů systému MIaS z akce MIR Happening (součást workshopu DML, konference CICM 2012).

Po typografické stránce je práce velice zdařilá. Je psaná v anglickém jazyce a její jazyková úroveň je dobrá.

Obsahově je práce na výborné úrovni. Text je napsaný velice konzistentně, jednotlivé části na sebe logicky navazují, autor pečlivě a smysluplně cituje. Veškeré své úvahy, závěry a návrhy řešení velmi dobře podkládá zjištěnými výsledky zkoumání.

Vysoce lze hodnotit druhou část kapitoly dvě – vyčerpávající rozbor možných přístupů k matematickému vyhledávání a možností využití klasických evaluačních IR metod. Stejně tak je možné hodnotit rozbor technik normalizace matematické formule v kapitole 4, kde autor přesvědčivě obhájí jím navrhovaná řešení.

Samotný software MIaS je dobře objektově navržen a plně funkční – jeho úspěšnost dokazuje i nasazení v EuDML. Z uživatelského hlediska je zadávání dotazů uděláno velmi pěkně i poměrně komfortně (dotaz se okamžitě při zadávání renderuje do výsledné rovnice – k tomuto účelu je použita knihovna třetí strany).

Evaluace systému MIaS je také provedena velmi dobře a dokazuje, že systém je funkční a dokáže zvládnout i poměrně komplikované dotazy, které navrhli přímo matematici. Zároveň ukazuje, kudy je možné systém ještě dále rozvíjet.

Poznámky a dotazy k obsahu práce:

- Kdo nakonec „vyhrál“ MIR Happening (viz strana 10)?
- Na straně 13 v popisu měření hodnocených výsledků je faktická chyba – v případě, že $(k + 1)$ výsledek je nerelevantní, pak přesnost klesne (a nikoli „increases“ jak je uvedeno).
- Na straně 27 se píše o neúspěšném zkoušení knihovny UCML? Proč se ukázala knihovna jako nevhodná?
- Jaké jsou šance, že se MlaS dostane přímo do Lucene? Proběhla už nějaká jednání s Apache Foundation?

Celkově hodnotím práci jako vynikající. Řešitel dokázal velmi dobře navázat na svoji bakalářskou práci a dále ji rozvinout. Zároveň dosáhl přínosných výsledků, které prezentoval (a bude prezentovat) i v zahraničí, a úspěšně je prakticky uplatnil v celoevropském projektu EuDML. Bez pochyb tak prokázal schopnost zvládnout danou problematiku a dále ji zajímavě rozvíjet. Celá práce je navíc zpracována velice pečlivě a kvalitně a splňuje zadání. Mohu prohlásit, že se jedná o jednu z nejzajímavějších a nejlepších absolventských prací, jaké jsem měl možnost dosud posuzovat. Práci hodnotím stupněm **A** a doporučuji ji navrhnout na Cenu děkana.

Obhajoby se pokusím zúčastnit osobně.

V Brně dne 4. února 2013.

Mgr. Vlastimil Krejčír
ÚVT MU