

# Posudek oponenta na diplomovou práci

## Transakce ve fulltextovém vyhledávacím stroji

Autor práce: Jakub Podhorný

Cílem práce je adaptovat transakční techniky pro práci s indexem fulltextového vyhledávače. Součástí je implementace nad platformou Java2, založená na vyhledávači Egothor.

Práce je členěna do velmi malých kapitol (běžná je délka 1-2 odstavce), jejichž číslování místy využívá až čtyři úrovně.

Kapitola 1 popisuje řešený problém a zároveň v jedné větě zavrhuje všechny existující implementace transakčních systémů jako nevyhovující.

Kapitola 2 vysvětluje, jak probíhá práce z indexem v systému Egothor. Je zaměřena nejen na popis samotné funkce, ale zároveň zabíhá – podle mého názoru zbytečně – i k implementaci (obsahuje hlavičky funkcí v Javě, vysvětluje význam některých tříd,...).

Kapitola 3 se převážně zabývá obecně transakčním zpracováním, zároveň však ukazuje jejich vztah k řešenému problému. Jsou popsány problémy, které nastávají při práci s transakcemi (deadlock, starvation,...), a existující přístupy k transakčnímu zpracování (timestamp ordering, certification, multiversion processing,...) spolu s diskusí o jejich vhodnosti pro použití ve fulltextovém vyhledávači. Na závěr je zvoleno řešení, které je kombinací zamykání a multiversion principu.

Kapitola 4 podrobně popisuje zvolené řešení. Je napsána z hlediska toho, jak systém řeší jednotlivé problémy, bez zbytečných implementačních detailů.

Příloha A.1 obsahuje náznak testů provedených s vytvořenou implementací. Jedná se však pouze o testy velmi malého rozsahu, které mají na malých datech ověřit funkčnost řešení.

Zcela chybí testy, které by se alespoň blížily reálnému provozu, případně alespoň vypočtený odhad toho, jaký vliv bude mít přidání transakcí na výkon celého systému. Na příloženém CD jsou pouze zdrojové soubory, bez popisu postupu, kterým aplikaci zkompileovat a zprovoznit. Není možné posoudit, nakolik jsou oba problémy zapříčiněné autorem práce a nakolik jsou „zdeděné“ ze systému Egothor.

Práce je psána v angličtině, bohužel obsahuje velké množství gramatických chyb a podivných formulací. Naštěstí tyto problémy v naprosté většině případů nebrání pochopení práce, takže k nim v hodnocení nebylo přihlíženo.

Jako jeden příklad za všechny: „In this chapter there has been used Egothor project documentation by RNDr. Leo Galamboš, Ph.D. [1] even, though, it mainly covers different parts of the project.“

V citacích chybí odkaz na BerkeleyDB.

**Závěr:** Samotný návrh řešení působí celkem promyšleně, bohužel výsledek není možné otestovat a ani práce chování implementace v reálném provozu neuvádí.

I přes chybějící testy práce podle mého názoru splňuje požadavky kladené na diplomovou práci a jako takovou ji doporučuji k obhájení.

V Praze dne 20.8.2007

Mgr. Jiří Dokulil

