

Posudek vedoucího/oponenta* diplomové práce

Jméno a příjmení autora posudku: David Bednárek

Jméno a příjmení autora práce: Lucie Klimperová

Název práce: **SQL front-end pro systém Bobox**

Vlastní text (sem prosím napište text posudku, délka textu posudku není omezena):

Tématem práce je překlad dotazů v jazyce SQL do rozšířené relační algebry. Původním cílem bylo napojení výstupu na již existující software realizující konverzi relační algebry do podoby exekučního plánu pro prostředí Bobox. Tento plán, tj. realizace kompletního SQL front-endu v rámci dvou diplomových prací, se však ukázal jako poněkud nerealistický. Ačkoliv software vytvořený v rámci této práce produkuje výstupy použitelné z hlediska formátu v následných fázích celého systému, jako celek je zatím systém použitelný jen pro velmi jednoduché dotazy - pro složitější případy obvykle chybí pokrytí v některé z návazných fází. Příčinou jsou rozdílné priority vývoje jednotlivých fází - zatímco v této diplomové práci byl kladen důraz na širí pokrytí konstrukcí SQL a jsou tedy realizovány i typové konverze a podobné méně frekventované konstrukce, návazný software z předchozí práce byl veden snahou řešit především optimalizaci a další náročné problémy. Tato práce se naopak jednomu těžkému problému vyhnula - není realizován převod korelovaných poddotazů do relační algebry.

Vzhledem k důrazu na pokrytí jazyka je zřejmé, že tato práce představuje především velké množství práce spojené se zpracováním řady konstrukcí jazyka SQL. Zároveň se těžiště prací dostává do oblasti syntaktické analýzy a typové kontroly, tj. do oboru překladačů, a vzdaluje se tak od problematiky databází - konverze do relační algebry sice není triviální, ale představuje spíše menšinu práce. I samotná syntaktická a typová analýza dotazů v SQL však překračuje rozměry běžné diplomové práce, proto je pokrytí tohoto jazyka řešeno ve dvou vrstvách: Syntaktická analýza, tj. konverze gramatiky do podoby vhodné pro LALR(1) analýzu, byla provedena pro celý jazyk dotazů, jak jej definuje SQL:1999. Další kroky, tj. primární vnitřní reprezentace, typová analýza a konverze do relační algebry, jsou implementovány pouze pro vybranou podmnožinu jazyka, která však je dostatečná pro pokrytí většiny dotazů poměrně komplikovaného benchmarku TPC-H, přičemž nepokryté dotazy lze ručně upravit do pokryté podoby.

Výše zmíněná neúplnost kompletního řetězce (za kterou autor této práce nenese odpovědnost) znamená, že výstupy z této práce zatím není možné prověřit spuštěním. Lze však ověřit konformanci vůči požadovanému (dříve definovanému) formátu a dále ručně prověřit správnost překladu, neboť výstup (relační algebra serializovaná do XML) je v zásadě čitelný. Na základě takto provedené namátkové kontroly lze konstatovat, že odevzdaný software je funkční, ikdyž se dá očekávat, že intenzivní testování by nějaké nedostatky odhalilo.

Vzhledem k tomu, že řešení této práce bylo spíše pracné než intelektuálně náročné, nelze v textu práce očekávat teoretická pojednání či algoritmy, spíše jde o zprávu popisující vývoj software. Přesto jsou v práci dvě témata, která zaslouží podrobnější popis. Prvním tématem jsou úpravy gramatiky do podoby vhodné pro analýzu nástrojem bison - ačkoliv jde o problém mnohokrát v minulosti řešený, nejde o mechanickou záležitost ale o pečlivou práci vyžadující systematický přístup a důkladné porozumění teorii LALR(1) analýzy. Druhým problémem je návrh vnitřní reprezentace - jde sice v zásadě o klasický problém objektového návrhu, ovšem ve velkém rozsahu několika desítek vzájemně provázaných tříd.

Oba tyto problémy jsou v textu práce podrobně rozebrány, bohužel však autorovi schází vzácná schopnost popisovat složité problémy srozumitelným způsobem - nejzajímavější části textu jsou tak pro čtenáře velmi obtížně čitelné až nesrozumitelné. Některé části pak nesou známky příliš kvapného dokončování a nevhodné typografie. Text je tedy použitelný jako referenční dokumentace pro případné pokračovatele, ale nelze jej doporučit jako učebnici pro někoho, kdo se podobné dílo chystá zopakovat.

Práce tedy v zásadě naplnila zadání - v softwarové části bez vážnějších výhrad, v textové části s výše uvedenými výhradami.

Doporučení k obhajobě:

Z výše uvedených důvodů práci *doporučuji* / *nedoporučuji** k obhajobě.

Vynikající práce vhodná pro soutěž studentských prací	ANO <input type="checkbox"/>
---	------------------------------

Seznam soutěží studentských prací, viz <http://www.mff.cuni.cz/studium/bcmgr/prace/>

Pokud jste výše zaškrtnli ANO, zdůvodněte prosím svůj návrh, případně uveďte konkrétní soutěž, pro kterou je práce vhodná (rámeček lze nechat prázdný, pokud za dostatečné zdůvodnění považujete text posudku):

--

V Praze dne: 28.8.2015

Podpis:**

** nehodící se škrtněte (vymažte)*

*** do SISu vkládejte formulář nepodepsaný (ve formátu PDF), podpis je potřeba doplnit až na vytištěný posudek.*