



Posudek oponenta závěrečné práce

Oponent práce: Mgr. Jiří Toušek
Student: Bc. Ondřej Hlaváč
Název práce: Analýza datových toků ve skriptech v SAP Hana dialektu SQL
Obor / specializace: Softwarové inženýrství
Vytvořeno dne: 25. května 2022

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Práce splnila zadání.

Implementované prototypové řešení je dostatečně kvalitní a splňuje všechny technické požadavky softwarového systému Manta, je tak možné jej přímo využít v praxi.

2. Písemná část práce

65 /100 (D)

Rozsah práce odpovídá řešenému problému. Práce je dobře strukturovaná, text je pochopitelný. Některé podkapitoly kapitoly 2 však patří spíše do kapitoly 1, jedná se o obecné koncepty, ne o popis technologií. Kapitoly 3.4.4.1 a 3.4.4.2 jsou špatně zanořené, měly by být o úroveň výš. Odkaz v kapitole 2.5.2 je rozbitý. Odkaz v kapitole 3.4 měl patrně vést jinam (možná na kapitolu 3.3.7?).

V kapitole 3.3.1 bych očekával detailnější popis zajímavých konstruktů, které SQLScript poskytuje oproti obdobným jazykům v jiných databázích. Požadavky FR4 a NFR2, uvedené v kapitole 3.1, nejsou v dalším textu nijak adresovány. V kapitole 3.3.5.1 tabulka 3.3 neuvádí logické binární operátory - nevím, zda je to záměr, v tabulce 3.2 logický operátor "NOT" uveden je. Ve stejné kapitole je zmíněna precedence operátorů, podrobnosti však analýza neobsahuje - bylo by vhodné uvést alespoň odkaz do dokumentace. V kapitole 3.3.5.2 je popsána jen jedna ze dvou variant výrazu CASE. Ocenil bych detailnější analýzu klauzule ORDER BY (3.4.1) - na jaké objekty odkazuje a zda lze použít číselné indexy místo jmen sloupců. Naopak znalost některých základních konceptů by snad bylo možné od čtenáře očekávat (např. graf definovaný v kapitole 1.1, unární a binární aritmetické operace popsané v kapitole 3.3.5.1).

Závažné věcné chyby jsem v práci nenašel. Některé drobnější nepřesnosti a logické chyby:

- v definici grafu v kapitole 1.1 je uvedeno "u and u are vertices" místo "u and v are vertices";
- v kapitole 3.2.2 autor uvádí, že ANTLR používá LL(*) algoritmus pro vygenerování parseru
- toto je sice technicky pravda, avšak nepříliš podstatné konstatování, důležitější je, že LL(*) algoritmus používá vygenerovaný parser;
- kapitola 3.3.5.4 uvádí, že agregační funkce lze použít pouze spolu s GROUP BY, přitom např. SELECT COUNT(1) FROM TBL je validní příkaz;
- v kapitole 3.3.7.1 je uvedena množinová operace UNITE / UNITE ALL, dokumentace ale uvádí UNION / UNION ALL;
- v kapitole 3.3.7.5 autor píše, že CREATE TABLE ... AS SELECT nekopíruje data, pokud není uvedeno WITH DATA, toto tvrzení je zopakováno v kapitole 3.4.4.2, avšak ve stejné kapitole je v ukázce kódu 3.21 WITH DATA vynecháno a ilustrace 3.10 ukazuje tok dat - z dokumentace se zdá, že CREATE TABLE ... AS SELECT ve výchozím chování kopíruje data, tedy správně je ilustrace;

Práce obsahuje určité množství jazykových chyb (např. "error-prone" použité ve smyslu "odolný vůči chybám", "compiles" místo "combines"), toto však není na překážku pochopitelnosti textu. K typografické stránce nemám až na stranu 47 výhrady.

3. Nepísemná část, přílohy

95 /100 (A)

Prototypová implementace dodržuje základní zásady psaní kvalitního kódu, je dostatečně čitelná a okomentovaná. Implementace dodržuje standardy a strukturu projektu Manta. Implementace je řádně pokryta testy.

4. Hodnocení výsledků, jejich využitelnost

90 /100 (A)

Výsledky práce jsou přímo uplatnitelné v praxi.

Vzhledem ke kvalitě prototypové implementace a dodržení technických požadavků systému Manta lze bez větších problémů implementovaný prototyp převzít jako základ pro praktické řešení.

Celkové hodnocení

80 /100 (B)

Práce řeší všechny podstatné aspekty analýzy datových toků v databázi SAP Hana.

Autor musel zvládnout složitou problematiku parsování SQL skriptů a tvorby datových toků.

Celkovou kvalitu práce bohužel poněkud snižuje písemná část, zde především postrádám rozsáhlejší diskuzi specifik analyzovaného SQL dialektu. Autor musel mnoho konstrukcí a chování specifických pro SAP Hana při implementaci prototypu řešit, je škoda, že se jejich popis nedostal do písemné části.

Implementované prototypové řešení je technicky na vysoké úrovni, dobře navržené a splňuje všechny na něj kladené požadavky. Bylo provedeno ověření korektnosti implementace.

Řešení je přímo využitelné v praxi.

Otázky k obhajobě

-

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.