

Posudek oponenta bakalářské práce

Student: Bašta Jiří
Téma: Generátor testovacích dat pro relační databáze (id 19230)
Oponent: Rychlý Marek, RNDr., Ph.D., UIFS FIT VUT

- Náročnost zadání** **průměrně obtížné zadání**
Práce popisuje návrh a implementaci aplikace pro sestavování šablon pro generování náhodných dat vkládaných do databáze a aplikace generující a vkládající data dle uvedených šablon. Zadání nabízí možnosti řešit řadu zajímavých problémů, které však studentem nebyly využity (např. pokročilé generování cizích klíčů, hodnot atributů s explicitními omezeními domén, aj.). V kontextu předloženého řešení lze považovat zadání za průměrně obtížné.
- Splnění požadavků zadání** **zadání splněno s drobnými výhradami**
Zadání je splněno s drobnou výhradou k bodů 2 a 4, kdy problematika efektivního vkládání velkého množství dat mohla být diskutována důkladněji (nikoliv pouze jednotlivý a hromadný insert) a implementace mohla podporovat více integritních omezení, jak je uvedeno dále v posudku v bodech 7 a 8.
- Rozsah technické zprávy** **je v obvyklém rozmezí**
Technická zpráva je svým rozsahem v obvyklém rozmezí (od úvodu po závěr má 38 vysázených stran).
- Prezentační úroveň předložené práce** **60 b. (D)**
Technická zpráva má logickou strukturu, která odpovídá procesu vývoje programového řešení. Kapitola 4 "Návrh generátoru" je příliš stručná (6 stran) a vlastní návrh aplikace, tedy její architektura, datový model, třídy, aj., v ní není vůbec popsán. Za zcela nevhodné považuji zařazení kap. 4.3 "Zvolené technologie" do popisu návrhu aplikace, protože uvedená kap. 4.3 popisuje technologie obecně, nikoliv v kontextu aplikace (a byla by tedy vhodná někde v úvodu technické zprávy). Algoritmy v kap. 5 jsou nevhodně prezentované útržky PHP kódu a jsou obtížně srozumitelné (komentář chybí).
- Formální úprava technické zprávy** **70 b. (C)**
Z hlediska formální úpravy obsahuje technická zpráva malé množství chyb v typografii (např. mezera před tečkou v kap. 2.7 a před odkazem poznámky pod čarou č. 3 str. 16), gramatické (chybná interpunkce v 2. souvětí kap. 3.1), či v pojmenování a číslování ("listing" 5.1 a 5.2 na str. 23-24). Chyby však nemají větší vliv na čitelnost práce.
- Práce s literaturou** **80 b. (B)**
Seznam literatury obsahuje 18 položek, z nichž všechny jsou v práci odkazovány. Použité zdroje velmi dobře pokrývají řešenou problematiku. Nevhodná je citace s odkazem poznámkou pod čarou na konci kap. 2.5, která měla být zařazena v seznamu literatury.
- Realizační výstup** **60 b. (D)**
Technické řešení má drobné nedostatky. V kap. 3.3 nebyly diskutovány (a dále použity) některé další metody optimalizace vkládání dat - např. použití jedné transakce, použití předpřipravených dotazů (tzv. "prepared statements"), odesílání SQL příkazů v dávkách, použití speciálních nástrojů (např. mysqlimport). V kap. 5.2.3 pak nebyla řešena explicitní integritní omezení, např. domény či složené klíče entitní a referenční integrity. Výše uvedené nedostatky se promítly i do kvality programového řešení, kde lze za velkou chybu považovat např. sestavování SQL příkazů insert v metodě Generator.StartGenerate konkatenací textových řetězců namísto použitím substituce (printf či PreparedStatement).
- Využitelnost výsledků**
Výsledky jsou využitelné v praxi pro testování aplikací nad generovanými daty, avšak s několika výhradami - realizovaný generátor neumožňuje generovat data se základními omezeními domén (např. při generování textových hodnot vlastních jmen či emailových adres není možné tyto omezit délkou), s hodnotami primárních či cizích klíčů složených z více atributů, s cizími klíči rekurzivních vztahů mezi tabulkami (např. vazba nadřizený-podřizený v tabulce zaměstnanci) aj.
- Otázky k obhajobě**
 - V kap. 4.1.3 zmiňujete, že v případě automatizované tvorby šablony "ne všechny její možnosti jsou ideální a bezpečné". Prosím zdůvodněte toto tvrzení, zejména zmiňovanou bezpečnost.
 - Na str. 17 v obr. 4.2 jsou zobrazeny výsledky průzkumu uživateli preferovaného způsobu vkládání dat. Kolik uživatelů se účastnilo průzkumu a jakým způsobem probíhal?
 - Součástí realizačního výstupu jsou soubory s rozsáhlými seznamy vlastních jmen a jiné. Odkud tato data pocházejí?

10. Souhrnné hodnocení

60 b. uspokojivě (D)

Vzhledem k drobným výhradám ke splnění zadání a nižší kvalitě technické zprávy a programového řešení navrhuji hodnotit práci na stupněm **uspokojivě (D)**.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 31. května 2017

.....
podpis