

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy v Praze

Autor práce David Kuboň
Název práce Framework pro extrakci informací z velkého množství jazykových dat
Rok odevzdání 2014
Studijní program Informatika **Studijní obor** Obecná informatika
Autor posudku David Bednárek **Role** Oponent
Pracoviště KSI MFF UK

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Komentář Řešená úloha je v principu velmi jednoduchá - na hranici přijatelnosti pro Bc. práci – při důkladném zohlednění požadavků na výkon však může být zkomplikována až na úroveň Mgr. práce. Požadavek na výkon byl explicitní součástí zadání, autor bohužel výkonové požadavky zohlednil fakticky pouze volbou jazyka C++. Výsledné dílo zřejmě splňuje požadavky na funkčnost, ve výkonnosti však zůstávají značné rezervy – viz níže.				

Textová část práce

	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Analýza	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komentář Jazyková i typografická úroveň vyhovuje až na kolokvialismy typu „mečování“. V práci chybí jednoznačný popis požadované funkčnosti, ani klíčové algoritmy (viz např. 3.5.2) nejsou dostatečně přesně popsány. Analýza je sice poměrně rozsáhlá, zcela však ignoruje sofistikovanější datové struktury, přestože na jiném místě práce cituje články, ve kterých jsou úspěšně použité datové struktury popsány.				

Implementační část práce

	lepší	OK	horší	nevyhovuje
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
<p>Komentář Autor zvolil jazyk C++ z důvodu rychlosti, program je však napsán stylem, jakým se programuje v jazyku Perl a příbuzných. Především jde o neustálé kopírování řetězců i složitějších datových struktur, dané zřejmě jednak neschopností použít obraty specifické pro C++ (např. iterátory), jednak zjevnou neznalostí základních zvyklostí v C++ (předávání řetězcových parametrů hodnotou). Extrémním případem jsou pak odstrašující datové struktury typu <code>vector<vector<vector<string>>></code>. Důsledkem je, že autorova implementace pravděpodobně nebude o nic rychlejší, než obdobná implementace v autorem kritizovaných „interpretovaných“ jazycích.</p> <p>Implementace je citlivá na správné umístění dat a nastavení konfigurace – některé chyby nejsou diagnostikovány a vedou k nečekaným výsledkům. Autor na tuto skutečnost v návodu upozorňuje, vzhledem k určení aplikace by však měla být diagnostika lepší, neboť uživatel typicky nebude schopen posoudit správnost výsledků.</p>				

Celkové hodnocení Dobře Choose an item.

Práci navrhuji na zvláštní ocenění Ne

Datum 9. června 2014

Podpis