

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Václav Motyka
Oponent práce: RNDr. Ondřej Suchý, Ph.D.
Název práce: Efektivní algoritmy pro řešení problému nalezení konvexní obálky v 3D
Obor: Teoretická informatika

Datum vytvoření: 12. 6. 2017

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
Komentář: Zadání považuji za mírně náročnější. Algoritmy jsou netriviální a geometrie v prostoru vyžaduje určitou míru obezřetnosti. Pro návrh optimalizací a paralelizace je pak potřeba algoritmy detailně pochopit.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
Komentář: Zadání z větší části nebylo splněno. Algoritmy dle části 1) sice asi nastudovány byly, jejich popis v práci však nesevřídčí o tom, že by student algoritmy hluboce pochopil včetně detailů nutných pro jejich správnou funkčnost. Z popisu v práci je těžké, ne-li nemožné, algoritmy pochopit. U geometrických algoritmů je často velký rozdíl mezi geometrickou intuicí a skutečnou manipulací s datovými strukturami. Práce však tyto dva pohledy ignoruje a pokud už jsou uvedeny, nijak je nepropojuje. Dle části 2) zadání byly implementovány pouze dva algoritmy, u třetího byla použita dostupná implementace a jen zaměněny celočíselné proměnné za proměnné realizující čísla s plovoucí desetinou čárkou. Analýza a návrh paralelizace v OpenMP (část 3) zadání se omezuje na jeden odstavec pro každý algoritmus, konstatující, které části algoritmů by mohly být vhodné k paralelizaci a jaké problémy by to přineslo. Pro generování bodů (část 4)) je použito volání existujícího programu s omezenými možnostmi. Generování bodů vlastním programem nebylo implementováno. Výkonost algoritmů byla skutečně porovnána na serveru Star, ale o ověření správnosti s existujícím řešením se práce nezmiňuje (část 5) zadání).	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
Komentář: Rozsah splňuje požadavky, i když některé sekce by mohly být považovány za zbytečné -- například Sekce 2.1.3, Sekce 3.5, Sekce 4.2 (viz poslední odstavec).	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

4. Věcná a logická úroveň práce

50 (E)

Popis kritéria:
Posuďte, 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.

Komentář:

Členění práce na kapitoly je obstojné, jen název kapitoly "Softwarové a hardwarové aspekty" je nevhodný, neboť jediná zmínka o hardwaru v celé práci je popis konfigurace serveru Star v kapitole Výsledky. Značná část kapitoly Optimalizace je věnována úpravám které jsou pro správné fungování algoritmu nutné.

V kapitole Výsledky chybí přehled toho, co bylo skutečně implementováno.

Řada definic je značně vágní a, pokud by čtenář nevěděl, co daný pojem znamená, těžko by to z definice zjistil. Definice mnohdy neodpovídají zvyklostem v oboru a na Fakultě informačních technologií. Pro formální definici mnohostěnu je čtenář odkázán do literatury, ačkoliv práce dříve zavedla pojem polytop s totožným významem. Podobně u definice kolineárních bodů není využita dřívější definice afinně závislých bodů.

V práci se často mluví o seřazení bodů po směru hodinových ručiček, což bez dalšího v prostoru dimenze 3 nedává žádný smysl. Pokud by se měl takový pojem používat, je potřeba ho definovat přesněji -- například "při pohledu z vnějšku polytopu". Ovšem ani pak není snadné takovou vlastnost bodů testovat. Přitom udržení správné orientace bodů je pro některé z představovaných algoritmů (např. QuickHull) zcela zásadní a je potřeba ho v implementaci značně ošetřovat. Podobně lehkovážně práce nakládá s pojmy "nad rovinou" a "pod rovinou", apod. Práce dále mluví o záporné vzdálenosti mezi bodem a rovinou, nebo o záporném objemu mnohostěnu.

Analýza složitosti Jarvis March ignoruje volání algoritmu pro 2D konvexní obálku, náznak analýzy složitosti Divide and Conguer algoritmu ignoruje to, že některé kroky jsou (potřeba) za konstantní čas jen amortizované. Uvedená složitost inkrementálního algoritmu je nesmyslná a (tudíž) analýza špatně. U randomizované verze se mluví o "předpokládané složitosti" -- zjevně špatný překlad anglického "expected complexity". U složitosti QuickHullu v průměrném případě není uvedeno v jakém smyslu je případ průměrný. Sekce 1.3 zaměňuje dolní mez složitosti algoritmu a problému. Velmi vágní je také popis "efektivity algoritmů" v sekci Cíl práce.

Rozpaky budí také formulace "reálných neboli celých čísel" v Instalační příručce. Přitom jinde se dočteme, že "naše algoritmy budou pracovat na tělese reálných čísel".

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

5. Formální úroveň práce

60 (D)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

Komentář:

Práce obsahuje řadu překlepů (i v abstraktu), které jsou zjevné i při letmém čtení.

Převzaté obrázky (tedy všechny kromě grafů výkonu) často nezapadají do textu a jsou nekvalitní. Ani jediné vlastní obrázky, tedy grafy výkonu, nejsou vloženy jako vektorové a jejich kvalita je nedostatečná.

Pro některé objekty se používá více různých symbolů.

Seznam zkratk uvádí nepoužité zkratky a neuvádí použité.

V sekci Struktura práce se mluví o algoritmech z "kapitoly dva" místo použití reference na kapitolu "3".

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

70 (C)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a uvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

U reference [28] chybí jeden z autorů. U reference [11] jsou jména některých autorů nahrazena zkratkou aj. (zřejmě dle ČSN 690), což je u práce z Teoretické informatiky značně nevhodné.

K některým tvrzením by se jistě dali najít relevantnější zdroje než wikipedia.

U jednoho z převzatých obrázků chybí odkaz na zdroj.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

40 (F)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvoril sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Zásadním bodem výstupu práce by měla být implementace algoritmů.

Nově implementovány byly pouze dva algoritmy místo tří. Zdrojové kódy jsou nevhodně členěny, například algoritmus pro 2D konvexní obálku najdeme v souboru structures.cpp.

Některé komentáře jsou v češtině, jiné v angličtině a budí dojem, že jsou psány různými autory. Některé části kódu se zjevně nepoužívají, jiné jsou nedodělané.

Některé podprocedury, které by mohli být využity v obou algoritmech jsou implementovány pro každý zvlášť a to různě.

Implementace navíc používá nevhodné datové struktury (vektor pro 3 souřadnice), které limitují efektivitu výsledného programu.

I porovnání výkonu jednotlivých algoritmů je poněkud nepřehledné.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

Komentář:

Vzhledem ke špatné kvalitě implementace nejsou výsledky využitelné.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě

Popis kritéria:

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).

Otázky:

Myslíte si, že kontejner vector představuje způsob uložení tří souřadnic daného bodu s optimálním výkonem?

Jak se při měření rychlosti jednotlivých algoritmů měřil čas a kolik měření se průměrovalo?

Proč byla použita zhruba dva roky stará verze překladače GCC?

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

40 (F)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nesmí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Zadání z větší části nebylo splněno. Algoritmy nebyly nastudovány do detailů, implementovány byly pouze dva, paralelizace nebyla provedena, jejich správnost nebyla ověřena s existujícím řešením. Generování bodů bylo pouze převzato.

Chybí přehled toho, co bylo skutečně implementováno.

Definice jsou vágní a používají se nesmyslné pojmy a formulace (např. se často mluví o seřazení bodů po směru hodinových ručiček). Vzhledem k tomu, že se jedná o práci z oboru Teoretické informatiky, nelze přehlédnout také chyby ve vyjádřeních týkajících se složitosti algoritmů nebo vyjádření týkající se optimálního výkonu univerzálních kontejnerů pro specifické použití.

Všechny obrázky kromě grafů času jednotlivých algoritmů jsou převzaté.

Kvalita výsledné implementace je i ze softwareového hlediska špatná.

Proto práci nelze doporučit k obhajobě.

Podpis oponenta práce: