

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Parallel graph algorithms for GPU
Jméno autora:	Radek Cichra
Typ práce:	bakalářská práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra matematiky
Oponent práce:	Jan Volec
Pracoviště oponenta práce:	Katedra matematiky, FJFI ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Pro splnění zadání práce bylo potřeba jednak nastudovat některé moderní paralelní grafové algoritmy, jednak framework TNL, a jednak základní principy heterogenního programování (CPU vs. GPU). Zadání práce proto považuji za dosti náročné.	

Splnění zadání	splněno s výhradami
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Body 1-3 zadaní považuji za splněné, ačkoliv, striktně vzato, pro jeden ze tří v bodě 3 zadaných algoritmických problémů (MST) funkční implementace chybí. To je rozhodně škoda, nicméně za zásadní problém to nepovažuji.	
Bod 4 je splněn jen částečně, a z mého pohledu jeho nejpodstatnější část – srovnání s jinými algoritmy / konkurenčními implementacemi – bohužel zcela chybí.	
Bod 5 z mého pohledu splněn nebyl (je dokumentace někde k dispozici?).	
Následující bude patrně otázkou vkusu, nicméně i tak musím zmínit, že práce by se z mého pohledu lépe četla, kdyby na místo obsažení nekompletních částí s několikrát zmíněnou omluvou, že z důvodu XY tato věc nefunguje / nestihla se, tyto části ve výsledné práci vůbec nebyly. Určitě považuji za správné v úvodu resp. v závěru toto zmínit, ale ve zbytku textu by mi přišlo vhodné se soustředit čistě na to “prezentované”, nikoliv na to “mohlo/mělo by zde být prezentováno”. Asi nejkřiklavější případ tohoto je celá kapitola 5 o minimálních kostrách.	

Zvolený postup řešení	vhodný s výhradami
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Z celkového pohledu byl zvolen určitě správný postup ve smyslu nejprve si teoreticky nastudovat algoritmus, dále ho precizně popsat v pseudokódu, ten poté naimplementovat v konkrétním programovacím jazyce / frameworku, následně implementaci otestovat a změřit její efektivitu na různých vstupech, které pak budou porovnány s konkurenčními algoritmy či implementacemi. Má hlavní výhrada vůči zvolenému postupu řešení je, že se řešily výše uvedené body pro 4 algoritmy společně, a proto nakonec z časových důvodů na některé z těchto bodů nezbyl čas, a tudíž v práci obsaženy nejsou.	

Odborná úroveň

podprůměrná

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Algoritmus FastSV, pochopil-li jsem článek Zhang-Azad-Hu správně, je v pseudokódu prezentován chybně, a stejná chyba mi přijde, že se vyskytuje i v následné C++ implementaci. Konkrétně mi jde o kroky 4-9 v "Algorithm 2" na straně 17. Podle mého názoru Zhang-Azad-Hu nejprve provedou pro všechny hrany grafu tzv. "stochastic hooking", a až je toto provedeno, provedou opět pro všechny hrany grafu tzv. "agressive hooking".

Často se v práci vyskytují pojmy, jejichž význam si čtenář musí domyslet; kanonickým příkladem je pojem "convergence", byť zde jsem si z kontextu význam (doufám, že správně) domyslel. U některých dalších pojmů, které se v práci používají bez sebemenšího vysvětlení, to však tak snadné nebylo, a ne vždy mi bylo jasné, zda se bez toho čtenář obejde, nebo si to před čtením dalšího nejdříve musí domyslet. Zde jako typického reprezentanta považuji "stochastic hooking" a "agressive hooking".

Definovat pro graf v angličtině "the size" jako počet vrcholů osobně považuji za dost nešťastnou volbu; typicky totiž "the size" značí počet hran a počet vrcholů se v anglické odborné literatuře pojmenovává "the order".

Formální a jazyková úroveň

podprůměrná

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce se z mého pohledu velmi špatně čte. Osobně jsem zastáncem psaní prací v angličtině, nicméně zde mi přijde, že to srozumitelnosti ublížilo příliš. Celkově se mi dere na mysl otázka, jak konkrétně se Darren Copeland, MSc. (Language advisor) podílel na přípravě odevzdané práce.

Jedním z hlavních problémů, které jsem při čtení práce měl, je špatné používání ne/určitých členů. Na řadě míst totiž člen zcela chybí, ač by tam být měl. Proto čtenář musí být vysoce pozorný, jestli se např. v dané chvíli hovoří o "the vertices" (tzn. všech vrcholech právě uvažovaného grafu), nebo o "vertices" (tzn. nějaké, v danou chvíli nespecifikované, a ne nutně vlastní podmnožině vrcholů právě uvažovaného grafu). Problémy s členy jsem našel prakticky na každé straně práce a v průměru jejich počet překračoval pět chybných použití na jedné stránce.

Řada použitých anglických obrátů mi přijde přinejmenším pro vědeckou angličtinu nestandardní, byť nejsem rodilý mluvčí a tak to nedovedu 100% posoudit. Nicméně, uvozovat definice grafových pojmů pro daný graf "Suppose there is an undirected graph G. (...)" namísto, např., "Let G be an undericted graph. (...)" nepovažuji za dobrou angličtinu; viz. "Chapter 1".

Z typografického hlediska – formát jednotlivých citací je zcela nejednotný (občas zkrácené názvy časopisů, občas ne; občas je uveden, v napříč prací různém formátu, i měsíc publikace a občas uveden není vůbec, atd.).

Výběr zdrojů, korektnost citací

výborné

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

Použité a citované zdroje v práci považuji za adekvátní, a nemám k nim žádné výhrady.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Své výše zmiňované výtky bych souhrnně shrnul tak, že si "autor utrhl příliš velký kus koláče naráz, který pak nestihl sníst". Možná by dávalo větší smysl nejprve pořádně zpracovat a sepsat všechny aspekty jednoho algoritmického problému (včetně srovnání s jinými implementacemi!), a až poté přejít na další problém, zbyl-li by čas, nicméně, "po bitvě umí být generálem každý" a "chybami se člověk učí". Rozhodně nemám pochyb, že autor na vypracování předložené práce vynaložil značné množství svého času a úsilí, i přesto však bohužel práci v aktuálně prezentované podobě nepovažuji za příliš kvalitní.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Otázky:

- 1) Byl algoritmus FastSV implementován správně; viz. předchozí část posudku "Odborná úroveň" ?
- 2) Jaký je rozdíl mezi "stochastic hooking" a "agressive hooking"? Je za volbou těchto názvů nějaká motivace ?
- 3) Jaké hodnoty ve vaší definici "Adjacency Matrix" nabývají prvky této matice v případě grafu bez vah ?
- 4) Co znamená "mutually reachable" ?
- 5) Porovnával jste nějak u variant pro MIS (MIS_base, MIS_lex3, MIS_opt3) krom času také kardinalitu příslušným algoritmem nalezené nezávislé množiny pro stejný vstup (u pravděpodobnostní varianty algoritmu by dávalo smysl uvažovat průměr výsledné kardinality napříč několika nezávislými běhy) ?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **D - uspokojivě**.

Datum: 21.8.2024

Podpis: 