

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Efektivní výpočty vlastností neorientovaných grafů
Jméno autora:	Jana Zelenková
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Petr Ryšavý
Pracoviště oponenta práce:	IDA, Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>Úkolem práce bylo porovnat efektivitu dostupných programů pro počítání vlastností prostých neorientovaných grafů. Studentka měla za úkol vybrat množinu těchto vlastností, následně pomocí vlastních skriptů měřit čas běhu na vygenerované množině grafů a určit vliv typu grafu a použitého hardware. Zadání neobsahuje žádné body, které by představovaly riziko, že je nepůjde uskutečnit, náročnost je daná pouze pracností danou velkým množstvím zkoumaných metod.</p>	

Splnění zadání	splněno s menšími 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>	
<p>V bakalářské práci jsou zvoleny pouze dva balíky, pro které studentka porovnávala časy běhu, a to SageMath knihovna pro jazyk python a Wolfram Mathematica. Přehled vlastností, které jsou počítány považují za zcela dostatečný, jde dohromady o přibližně 12 vlastností zahrnujících spojitost grafu, hledání Eulerovských a Hamiltonovských kružnic, určování poloměru grafu a jiné. Naopak diskusi charakteristických vlastností považují za nedostatečnou, studentka rozděluje grafy pouze na spojitě, cyklické a náhodné. Očekával bych například porovnání hustých a řídkých grafů, především u problémů typu hledání minimální kostry či hledání Hamiltonovské kružnice. Porovnání efektivnosti při využití superpočítače obsahuje pouze sdělení, že výkonnější stroj nemá na výsledky vliv, protože uvažované programy jsou spouštěny jednovláknově. Porovnání výsledků s teoretickými předpoklady je provedeno pouze neformálně, vždy pár vět pro každou vlastnost. Daný přehled většinou neobsahuje ani očekávanou časovou náročnost problému.</p>	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
<p>Studentka zvolila správný postup řešení. Pro evaluaci používá kombinaci kódů v jazycích Bash, Python, Wolfram a C++. Kód je dokumentovaný, přehledný a rozšířitelný.</p>	

Odborná úroveň	D - uspokojivě
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
<p>Zde mám největší výhrady. Přehled grafových vlastností je pouze neformální, např. definice grafu (<i>O grafu můžeme přemýšlet jako o množině bodů, které jsou spojeny čarami, které vedou buď od jednoho bodu k druhému nebo z daného bodu zpět. Tyto body můžeme definovat jako vrcholy grafu a čáry mezi nimi jako hrany.</i>) Definice na sebe často nenavazují, například:</p> <ul style="list-style-type: none"> v kapitole 2.2 je zaváděn souvislý graf jako graf, jehož každé dva vrcholy jsou spojené cestou. Nikde předem není řečeno, co je cesta. Stejně tak definice stromu v kapitole 2.3 se opírá o chybějící definici cesty/kružnice. 	

- Ve 2.4 je definovaná excentricita vrcholu, přičemž se tato definice opírá o nikde nevysvětlenou vzdálenost, ta je až v 2.19.
- Ve 2.13 u vysvětlení chromatického čísla chybí informace o tom co je to diskrétní či bipartitní graf.
- V kapitole 2.14 se mluví o planárních grafech. Rovinné grafy jsou vysvětlené až v 2.16.

Některé neformální definice jsou nejednoznačné. Například popis NP-úplné úlohy (kapitola 3.3) sice ve dvou poskytnutých větách neobsahuje nic nepravdivého, ale autorka nevysvětluje, co to vlastně NP-úloha je nebo proč jde o důležitou množinu problémů. Po přečtení by pak uživatel mohl nabýt dojmu, že NP-úplné úlohy jsou obecně všechny, u kterých není znám polynomiální algoritmus.

Formální a jazyková úroveň, rozsah práce

B - velmi dobře

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

Text je místy špatně srozumitelný. Příkladem může být vůbec první odstavec bakalářské práce či věta v kapitole 2.8 (Eulerovský graf), kde studentka píše: „*Je překvapivé, že řešení tohoto problému je poměrně rychlé a jedná se o polynomiální nebo i lineární obtížnost.*“

V práci je malé množství překlepů, např. drobné chyby v gramatice – malé písmeno v první větě kapitoly 2, na straně 7 "Hamiltonovská graf". Množství měřených dat je prezentováno ve formě velkého množství grafů, vždy jeden graf pro každou vlastnost a každý program. U těchto grafů je navíc nevhodný popisek vodorovné osy – nejspíše mělo jít o počet hran, a ne počet vrcholů, který podle legendy grafu odlišen barvou. Z typografického hlediska bych vytknul pouze velké množství kapitol obsahujících často jen několik vět.

Výběr zdrojů, korektnost citací

C - dobře

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.

Určité výhrady mám jak k volbě zdrojů, tak ke způsobu citování. Studentka v přehledu důležitých grafových vlastností cituje vždy plošně dvě knihy představující přehled teorie grafů a k nim vždy zvolenou stránku z online dokumentace poskytnuté společností Wolfram na adrese <https://mathworld.wolfram.com/>. Jde sice o zdroj s uvedeným autorstvím a studentka navíc správně uvádí datum navštívení stránek, nicméně některé věty v textu v rešeršní části jsou pouhým doslovným překladem těchto stránek. Vzhledem k rozsahu popisů jednotlivých vlastností grafů, jde o nezanedbatelnou část textu, která by měla být označena jako doslovná citace. Například:

- *Jinými slovy, lze říct že průměr grafu je nejvyšší počet vrcholu, které musí být navštíveny pokud cestujeme z jednoho vrcholu do jiného, za předpokladu, že jsou vyloučeny cesty, které vedou zpět, jsou objížďkami nebo smyčkami. Proto se to rovná maximu všech hodnot v matici grafové vzdálenosti.* [kapitola 2.5]
- *In other words, a graph's diameter is the largest number of vertices which must be traversed in order to travel from one vertex to another when paths which backtrack, detour, or loop are excluded from consideration. It is therefore equal to the maximum of all values in the graph distance matrix.* [<https://mathworld.wolfram.com/GraphDiameter.html>]

Porovnání s citovanými knihami jsem neprováděl.

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.

Nemám další komentář. Studentkou vytvořené řešení je funkční, avšak prezentace ve formě velkého množství grafů je nevhodná. Očekával bych například navíc agregované výsledky, které by ukazovaly tabulku s porovnáním, na kterých problémech byla rychlejší knihovna SageMath, a na kterých Wolfram Mathematica.

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.

Studentka měla za úkol provést porovnání efektivity dostupných knihoven při výpočtu vlastností prostých neorientovaných grafů. Studentka si vybrala dvě knihovny, a to SageMath a Wolfram Mathematica. Dále měla porovnat vliv typu grafu a použitého hardware na čas výpočtu. Šlo o pracné, ale ne příliš náročné zadání.

Studentka zadání splnila, nicméně s určitými výhradami. Například vliv použití superpočítače na čas běhu byl označen za zanedbatelný s tím, že byly použity jedno-vláknové knihovny. Rozdělení grafů na spojitě, cyklické a náhodné také není příliš vhodné a diskuse této části je také příliš stručná. Práce obsahuje velmi málo formálních definic, přehled teorie grafů je neformální, z popisu NP-úplných úloh čtenář může nabýt dojmu, že jde z definice o úlohy, které nemají polynomiální řešení. Práce je místy špatně srozumitelná, ne příliš velkého rozsahu textu, některé myšlenky na sebe příliš nenavazují, kapitoly jsou krátké, definice využívají pojmů, které nejsou definované vůbec nebo až v následujících kapitolách. Za nevhodné považuji citace pouze tří zdrojů, a to dvou knih shrnujících teorii grafů a online dokumentace <https://mathworld.wolfram.com/>. Na druhou stranu kód, který studentka odevzdala je čitelný, s komentáři a rozšiřitelný.

Na studentku bych měl následující otázky.

- Jak byste si vysvětlila tvrzení na straně 11: *Zatímco Mathematica (viz graf 5.3) nezávisle na velikosti grafu, zjišťovala tuto vlastnost (myšleno, zda je graf stromem) v konstantním čase,* Je možné, že Mathematica některé vlastnosti vypočte již při načítání, které není zahrnuto do času běhu? Zahrnula jste do srovnání grafy, které mají stejný počet hran jako strom, ale přitom nejsou stromy, tj. s $m = n - 1$?¹ Jak je to v případě tvrzení *Zatímco Mathematica zvládá tuto vlastnost vypočítat v konstantním čase nezávisle na velikosti grafu u Eulerovských tahů* (strana 13)?
- V kapitole 5.3 mluvíte o strojích, které mají k dispozici okolo 1TB operační paměti. Měla jste k dispozici celou kapacitu stroje?
- Obrázky 5.21 a 5.22 ukazují jeden bod, který neodpovídá trendu. Byl odpovídající graf něčím výjimečný nebo jde z nějakého důvodu o outliera?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Datum: 1.6.2020

Podpis:

¹ m je počet hran grafu, n počet vrcholů