

Posudek vedoucího bakalářské práce

Autor: Karolína Rezková

Název: Rozšiřování knihovny pro grafové algoritmy

Posudek vypracoval RNDr. Marko Genyk-Berezovskyj, Katedra kybernetiky, FEL ČVUT

Autorka implementovala v knihovně JGraphT testování rovinnosti obyčejného grafu pomocí Hopcroft-Tarjanova Algoritmu. V porovnání s jinými grafovými algoritmy probíranými na FEL je to výrazně pokročilejší algoritmus, který se neprobírá ani v doktorském studiu. K implementaci, která je funkční a podle očekávání rychlá, podala několikastránkový poměrně přehledný popis hlavních myšlenek algoritmu kompilovaný z několika klasických zdrojů a doprovodila ho stručným popisem své implementace. V hlavních obrysech tedy zadání BP splnila. Nevěnovala se ale systematickému testování své implementace, jejíž spolehlivost proto zůstává neznámá, programovou dokumentaci vytvořila jen náznakově a hotovou práci po sobě nekontrolovala.

Tištěná část práce působí přehledným a úpravným dojmem. Autorka se snadno vyjadřuje a její text plyne velmi přirozeně. O to více při pozornějším zkoumání překvapí množství nedodělků, překlepů i chyb vyloženě věcných. Je patrné, že text práce po sobě autorka nečetla, a je to velmi nepříjemně znát. Ve výčtu implementovaných metod nacházíme nebývale zkomolená jména autorů rozšířených algoritmů „Floyd Warshallův“, „Tarianův“, „Edmont Karpův“, v Kuratowského větě, která je významná pro zkoumání planarity, je graf K_5 zaměněn za graf K_6 , čímž se smysl věty ztrácí, v úvodní kapitole najdeme definice ve zpřeházeném pořadí využívající pojmy definované až později, v jediném výpisu kódu v práci na str. 33 je zjevná syntaktická chyba apod.

Autorka ve své práci tvrdí, že testovala algoritmy již implementované i algoritmy, které sama kódovala a do knihovny přidala. Ke své práci ale nepřiložila žádný kód, který by toto testování skutečně prováděl. Také nepřiložila žádná data získaná z těchto testování, pouze v kap. 3.1 předložila dva sumární obrázky jen vágně popsané (a s překlepy v nadpisu). Nevíme tedy, co vlastně autorka dělala, jaká získala data a proč je interpretuje právě takto. Implementované testování rovinnosti selhává na stroji vedoucího práce při zpracování grafu s cca 7500 hranami. Objevuje se výjimka způsobená nedostatkem paměti JVM v podpůrné metodě (nikoli autorčině) hledající bloky v grafu. Přitom je k dispozici cca 1GB paměti, což by mělo bohatě při lineárním algoritmu stačit i na mnohonásobně větší graf.

Argumentace, že testovat grafy s více než 6 (šesti!) vrcholy nebylo možné pro jejich velký počet (str. 35), se nezakládá na faktech. Např. všech neorientovaných grafů se 7 vrcholy je právě 1044, nikoli 2 miliony, jak autorka mylně uvádí, a všechny jsou volně k přístupné na Internetu k experimentování (<http://users.cecs.anu.edu.au/~bdm/data/graphs.html>). Navíc, mezi grafy se 7 vrcholy je celkem pouze 220 grafů nerovinných. Dodatečné sestavení testovacího kódu a jeho spuštění na všech cca 12 milionech neorientovaných grafů s nejvýše 10 vrcholy zabralo vedoucímu práce necelou čtvrt hodinku.

V závěru tištěné práce autorka uvádí, že svůj kód doprovodila stručnou dokumentací v Javadoc-u. To je velmi diplomatické tvrzení. Co nazývá dokumentací, sestává celkem z 47 krátkých řádků převážně jen opisujících jména „getterů“ a „setterů“ a nedodržujících běžnou interpunkci, z toho dva řádky jsou zapomenuté české, u poměrně významné funkce `isGraphPlanar()` není dokumentován ani návratový typ ani návratová hodnota. U hlavního autorčina příspěvku – implementace Hopcroft-Tarjanova algoritmu – není uvedeno, podle kterého zdroje (původního článku?) byl kód sestaven, zda byly učiněny nějaké úpravy, optimalizace apod. Autorčin kód algoritmu používá další datové struktury, které nejsou zmíněny v tištěné části práce, jejich smysl není bezprostředně jasný.

V kódu funkce `isGraphPlanar()` se objevuje příkaz
`if (this.graph.vertexSet().size()<4) return true;`.
Zřejmě testovaná konstanta má být (alespoň) 5 a ne 4.

Autorka během vývoje své práce s vedoucím nekonzultovala, práci, stejně jako předcházející softwarový projekt, odevzdala v posledním možném termínu. Neformálně bych řekl, že spoléhá především na své nadání a zkušenosti ze studia na MFF UK, ke kterým ale nepřidává nezbytnou dlouhodobou vytrvalost a pozornost vůči detailům. Předkládaná práce tomuto stylu přesně odpovídá a projevuje se v ní zřetelně jeho kladný i záporný vliv.

Předloženou bakalářskou práci doporučuji k obhajobě a hodnotím známkou: **C – dobře**.

V Praze 12.6.2017

RNDr. Marko Genyk-Berezovskij