

## Posudek oponenta na diplomovou práci: Dotazovací engine grafové databáze využívající stromové dekompozice

**Autor práce: Bc. Karel Horák**

Předložená diplomová práce má tři cíle: Provést rešerši literatury týkající se homomorfismů grafů; vytvořit prototyp dotazovacího stroje pro grafovou databázi, který využívá stromovou dekompozici datazu pro efektivní vyhodnocování; a empiricky ověřit výkonost dotazovacího stroje na vybraných datasetech. V diplomové práci jsou všechny tři cíle splněny a to výborným způsobem.

Práce se věnuje grafovým databázím, které se stávají alternativou k tradičním relačním databázím. Jedná se tedy o tematiku, která je v současné době moderní a pro budoucí využití v praxi velmi důležitá.

Práce je rozdělena do šesti kapitol; po krátké úvodní kapitole následuje rešeršní část o grafových homomorfismech a grafových databázích. Ve třetí kapitole se čtenář seznámí se základními pojmy a postupy relačních databází. Nejdůležitější částí práce je čtvrtá kapitola (také svým rozsahem největší), která přináší popis nového algoritmu/dotazovacího stroje pro obecnou grafovou databázi. Autor se tu věnuje hlavně popisu algoritmu a důkazu jeho správnosti. Dále je uvedena i vlastní implementace a podrobně jsou popsány datové struktury, které algoritmus využívá. Velký prostor autor věnuje odvození a důkazu časové složitosti navrženého algoritmu. Na závěru kapitoly jsou ještě uvedeny tři algoritmy — první pouze rozhoduje, zda nějaký homomorfismus existuje, druhý zjišťuje počet homomorfismů a třetí jako výstup uvádí všechny homomorfismy daného grafu  $G$  do grafu  $H$ . Pátá kapitola přináší porovnání výsledků autorova algoritmu s výsledky Neo4j databáze. Testování je prováděno na různých datasetech; jedná se jak o „teoretické“ datasety, tak datasety generované z Barabasi-Albert modelu. Z výsledků testování je zřejmé, že nový algoritmus pracuje lépe než Neo4j. Poslední kapitola přináší shrnutí dosažených výsledků a možnosti pokračování výzkumu v dané oblasti.

Práce je velmi srozumitelně napsána a výborně se čte. Přestože autor všechny použité pojmy formálně definuje, též jednotlivé definice na příložených příkladech vysvětlu, což čtenáři velmi pomáhá. Práce s literaturou je také na výborné úrovni. Práce je napsaná velmi dobrou angličtinou. V práci jsem našla jen velmi málo překlepů.

K práci mám tyto drobné připomínky:

- V definici podgrafu na str. 6 by bylo vhodné zmínit, že podgraf musí sám též tvořit graf. Též by stálo za úvahu, zda mluvit o „regulárním“ grafu tam, kde je míněn graf.
- Při „sčítání“ tříd složitosti by autor měl být trochu podrobnější/přesnější. Bylo by dobré, kdyby autor v rámci obhajoby zdůvodnil následující rovnost, kterou uvádí v důkazu Věty 4.3. na str. 51

$$\sum_{i=1}^{|\chi(r)|} \mathcal{O}(i \cdot |V(\mathbf{H})|^i) \text{ což je } \mathcal{O}(|\chi(r)| \cdot |V(\mathbf{H})|^{|\chi(r)|}).$$

Práci považuji za výnikající, obsahem i vypracováním přesahující úroveň obvyklou pro diplomové práce. Její výsledky by se jistě daly publikovat v časopise či na konferenci.

Práci doporučuji k obhajobě a navrhuji známku A – výborně.

V Praze, dne 23. 5. 2015

Prof. RNDr. Marie Demlová, CSc.  
katedra matematiky 13101