

## Review of Bachelor's Thesis

**Student:** Šilling Petr  
**Title:** Applying Code Change Patterns during Analysis of Program Equivalence (id 24037)  
**Reviewer:** Fiedor Tomáš, Ing., Ph.D., DITS FIT BUT

- 1. Assignment complexity** **more demanding assignment**  
Práci považuji za obtížnější, jelikož vyžaduje (1) rozšíření netriviální techniky pro analýzu sémantické ekvivalence programů, a (2) efektivní implementaci umožňující analýzu projektů o velikosti linuxového jádra.
- 2. Completeness of assignment requirements** **assignment fulfilled**  
Zadání bylo splněno bez výhrad.
- 3. Length of technical report** **in usual extent**  
Vysázená práce čítá 42 stran (tzn. přibližně 63 normostran), čímž je v obvyklém rozmezí. Práce neobsahuje vatu; veškeré její součásti jsou relevantní a teoretický základ je dostatečně pokryt.
- 4. Presentation level of technical report** **90 p. (A)**  
Prezentační úroveň práce je kvalitní. Práce je psána v anglickém jazyce, čte se plynule (byť místy ji narušují trochu chaotické věty---např. začátek kapitoly 5).  
  
Drobné výhrady mám k (1) místy nedefinovaným proměnným v algoritmech (např.  $P_i$  a  $G_i$  v Alg. 2.1, nebo  $i_c^a b$  v Alg. 5.2), a (2) místy chybějící intuici (např. u Obrázku 3.2) nebo ilustrující příklady (např. SPCP, nebo běh hlavních algoritmů by zasloužily lepší ilustraci). Algoritmus 5.4 je dle mého názoru poměrně triviální a v práci by být nemusel.
- 5. Formal aspects of technical report** **90 p. (A)**  
Práce je z gramatického hlediska velice kvalitní. Student však (trochu ironicky) občas používá ne úplně sémanticky vhodné slova (např. "analysis revealed patterns") a ne vždy je konsistentní z hlediska britského/amerického dialektu (analyses vs analyzing). Rozšířený abstrakt obsahuje několik překlepů nebo nevhodných slov (např. majoritní verze).
- 6. Literature usage** **100 p. (A)**  
Práce cituje 22 zdrojů (primárně článků); nejsem si vědom porušení citační etiky. Student pokryl dostatečný vzorek relevantních zdrojů a prokázal jejich pochopení (oblast sémantické ekvivalence programů je rozsáhlá).
- 7. Implementation results** **95 p. (A)**  
Práce je realizována jako rozšíření nástroje diffkemp a jeho efektivita je demonstrována na porovnání několika dvojic jader kernelu. Vytvořené řešení je efektivní: přináší podstatné rozšíření nástroje za cenu minimální režie (dokonce nástroj diffkemp i mírně zrychluje). Práci jde zkompileovat, a experimenty reprodukovat. Výtku bych však měl k chybějící diskusi vlivu studentova řešení na míru falešných pozitiv.
- 8. Utilizability of results**  
Práce je používána v rámci nástroje diffkemp, který se používá ve firmě RedHat pro analýzu jader linuxu. Současně však vidím obrovský potenciál použití této verze diffkempu i pro další použití: např. pro optimalizaci výkonnostního profilování nebo jiných oblastí testování, jako je např. testování pokrytí. Práce byla rovněž prezentována na studentské konferenci Excel@FIT.
- 9. Questions for defence**
  1. Jak je řešeno pořadí aplikování šablon? Je aplikování šablon nezávislé na pořadí?
  2. Jak jsou řešeny nejednoznačné (*ambiguous*) šablony? Tzn. pokud pro danou instrukci lze aplikovat dvě různé šablony (např. díky kovarianci/kontravarianci typu operandů), která se použije?
  3. Diskutujte vliv Vašeho řešení na míru falešných pozitiv (nejlépe na experimentech uvedených v práci).
- 10. Total assessment** **95 p. excellent (A)**  
Práci Petra Šillinga hodnotím jako vynikající: (1) text práce je nadprůměrný, (2) vytvořené řešení je schopno analyzovat linuxové jádro, a hlavně (3) výsledek je použitelný v praxi.

Současně bych rád doporučil tuto práci k ocenění.

In Brno 30 May 2021

Fiedor Tomáš, Ing., Ph.D.  
reviewer