

Posudek oponenta diplomové práce

Student: Bančák Michal, Bc.
Téma: Nástroj na vizualizaci plagiátů v různých programovacích jazycích (id 18415)
Oponent: Burget Radek, Ing., Ph.D., UIFS FIT VUT

- Náročnost zadání** **průměrně obtížné zadání**
Zadání předpokládá zvládnutí existujícího poměrně komplexního nástroje pro generování syntaktických analyzátorů. Současně je však možno využít existujících, předem připravených analyzátorů a náročnost zadání proto hodnotím jako průměrnou.
- Splnění požadavků zadání** **zadání splněno s drobnými výhradami**
Všechny body zadání byly splněny. Klíčový nástroj ANTLR zmiňovaný v bodě 1 zadání je však popsán spíše okrajově v kapitole 3.1.1 a jeho možnosti a způsob práce nejsou zcela objasněny.
- Rozsah technické zprávy** **splňuje pouze minimální požadavky**
Rozsah technické zprávy nedosahuje dolní hranice obvyklého rozmezí, zcela jistě však splňuje minimální požadavky. Zejména vlastní návrh řešení mohl být popsán podrobněji stejně jako výše zmíněný nástroj ANTLR.
- Prezentační úroveň předložené práce** **67 b. (D)**
Technická zpráva poměrně pěkně shrnuje problematiku plagiátorství programového kódu a poskytuje stručný přehled existujících nástrojů v této oblasti. Základní metoda detekce plagiátů je dobře zvolená a návrh celého řešení je srozumitelný, ačkoliv způsob reprezentace abstraktního syntaktického stromu (AST) nepovažuji z hlediska další implementace za úplně šťastný. Popis implementace mohl být více strukturovaný; obsahuje dlouhé pasáže textu s různými implementačními detaily a podle mého názoru zbytečnými ukázkami kódu (např. str. 22), takže ve výsledku je poměrně nepřehledný a celkový způsob řešení v něm zaniká. Vyskytují se i drobné formální i faktické chyby v popiscích obrázků, názvech metod apod. Zvolený způsob reprezentace AST není nijak zdůvodněn a podle mého názoru je velmi nepraktický a komplikuje vlastní implementaci. Naopak testování je zpracováno poměrně přehledně.
- Formální úprava technické zprávy** **74 b. (C)**
Z typografického i jazykového pohledu je práce standardní pouze s občasnými drobnými prohřešky.
- Práce s literaturou** **80 b. (B)**
Seznam použité literatury je poměrně rozsáhlý a kromě zdrojů pokrývajících použité technologie obsahuje i vědecké články z oblasti detekce podobnosti zdrojových kódů. V textu práce jsou zdroje řádně citovány.
- Realizační výstup** **80 b. (B)**
Implementačním výstupem je plně funkční aplikace pro vyhledání plagiátů ve zdrojových kódech. Aplikace disponuje pěkným uživatelským rozhraním a je snadno použitelná. Kladně hodnotím i možnost grafického zobrazení syntaktických stromů. Ke způsobu implementace mám řadu výhrad - zejména kvůli komplikované objektové reprezentaci syntaktického stromu která často vede na velmi redundantní kód. Na druhou stranu však kód nese známky pečlivé práce a využití nástroje ANTLR je také dobře zvládnuté.
- Využitelnost výsledků**
Vytvořená aplikace umožňuje snadné porovnání dvou i více zdrojových souborů na případné plagiáty, z tohoto pohledu je dobře prakticky využitelná.
- Otázky k obhajobě**
 - Na obr. 4.4 uvádíte ukázkou kódu s mnohonásobným větvením programu podle typu zpracovávaného uzlu. Bylo by možné v jazyce Java vymyslet nějaké elegantnější a efektivnější řešení?
- Souhrnné hodnocení** **72 b. dobře (C)**
Pan Bančák v rámci své diplomové práce navrhl a vytvořil funkční a dobře použitelný nástroj pro detekci plagiátů pro různé programovací jazyky. K obsahu technické zprávy a některým částem implementace mám drobné výhrady, na druhou stranu celkově je implementace provedena pečlivě a metoda porovnání vstupních souborů je dobře zvolena a použita.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 6. června 2019

Burget Radek, Ing., Ph.D.
oponent