



Posudek oponenta závěrečné práce

Oponent práce: Ing. Miroslav Balík, Ph.D.
Student: Michal Dvořák
Název práce: Vyhledávání opsaných programů
Obor / specializace: Teoretická informatika
Vytvořeno dne: 25. května 2022

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

2. Písemná část práce 95 /100 (A)

Rozsah práce, návaznosti jednotlivých kapitol a i obsah práce jsou na vysoké úrovni. Drobnou výhradu bych měl k nečitelnosti obrázku na straně 17, attrMismatch označuje ohodnocení stavu, ve kterém se shodují typy tokenů, ale jejich atributy NE, v práci je chybně uvedeno, že ano. Vzorec na straně 22 by si zasloužil lepší vysvětlení, z textu není zřejmé, že se jedná o míru odlišnosti, nikoli podobnosti.

3. Nepísemná část, přílohy 100 /100 (A)

Příložený program je napsaný čitelně, všechny převzaté části jsou řádně odděleny.

4. Hodnocení výsledků, jejich využitelnost 100 /100 (A)

Výsledný program je nasaditelný do Progtestu.

Celkové hodnocení 95 /100 (A)

Pan Dvořák zpracovával zadání, které cílilo na vyhledávání plagiátů ve zdrojových kódech programů. Hezky zpracoval existující řešení a s ohledem na potřeby FITu se snažil vybrat takové, které by efektivně vyhledávalo ve zdrojových programech v C a C++ a to v objemech tisíců odevzdání úloh a miliónech dvojic pro porovnávání. Jeho řešení je odolné

proti prohození funkcí - jednomu z častých pokusů o maskování plagiátorství.
Práci doporučuji k obhajobě a hodnotím klasifikačním stupněm A - výborně.

Otázky k obhajobě

Vámi popsané algoritmy přistupují k porovnávání jednotlivých tokenizovaných odevzdání bez jejich dalšího předzpracování pro efektivnější vyhledávání. Realizujete tedy porovnání dvou odevzdání nezávisle na ostatních. Pro počty odevzdání v řádech desítek je to asi zbytečné, ale pro tisíce odevzdání by předzpracování mohlo být výrazně urychlující. Posun by se dal přirovnat k posunu od LCS ke K-LCS, tedy LCS pro množinu řetězců. Uvažoval jste o využití nějaké formy předzpracování řetězců tokenů pro efektivnější vyhledávání shod a podobností?

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.