



Posudek oponenta závěrečné práce

Student: Bc. Tomáš Dejmek
Oponent práce: Ing. Tomáš Pecka
Název práce: Finite tree automata to regular tree expressions conversion by removal of states
Obor: Systémové programování

Datum vytvoření: 3. 6. 2019

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	<u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
<p><i>Popis kritéria:</i> Posuďte, 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, posuďte 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.</p> <p><i>Komentář:</i> Práce měla za úkol představit regulární stromové výrazy (RTE) a konečné stromové automaty (FTA) a vymyslet nový algoritmus pro převod FTA na RTE, který bude inspirovaný metodou eliminace stavů konečného (řetězcového) automatu. Tento algoritmus bude implementován do ALT.</p> <p>Student měl za úkol vymyslet netriviální algoritmus, z toho důvodu považuji zadání považuji za náročnější.</p>	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	50 (E)
<p><i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, 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.</p>	

Komentář:

Rozsah práce je na spodní hranici doporučené délky (49 stran). Všechny části jsou ale relevantní a logicky navazují.

Práce je psána v angličtině a obsahuje větší množství gramatických a jazykových chyb (členy, časy, překlepy), které ztěžují pochopení textu. Je psána stylem, který je napůl beletristický a napůl sonda do aktuálních myšlenek autora ("I have to mull over", "to cut long story short", "the question comes to mind", ...).

V práci je plno doménově zvláštních slovních spojení, které bych si možná dovedl představit při neformálním seznámení s tématem, ale nikoli v závěrečné práci. Např. "trees are flowing into states", "transitions are weird edges", "ranked alphabet powered by special features" a další.

V práci se vyskytuje několik srovnání, underflow i overflow. Obrázky a algoritmy v plovoucích prostředích nejsou (až na jednu výjimku) odkazovány z textu, ale jsou umísťovány tak, aby spolu s popisem byly součástí textu. Některé obrázky ani nejsou v plovoucím prostředí (p. 18+19, 25). Definice 1.3.1 a 1.3.2 by měly být lépe strukturovány, např. pomocí seznamu s odrážkami. V Def. 1.3.4 chybí prostředí cases. Matematická sazba je občas zalomena na zvláštních místech, občas zalomena není a část řádky zůstane prázdná (p. 9, 38). Kolem matematiky je občas zvláštně umístěná interpunkce (p. 18 a 23). Názvy definic jsou nekonzistentní ve velkých písmenech. V sekcích 5.2.1 a 5.2.2 se vyskytují verbatim prostředí, které by měly být lépe odkazovány z textu, neboť takto není pro čtenáře jasné, o co se jedná. V kapitole 5 by mělo být ve shrnujícím prvním odstavci více textu. V sekci 5.4 není jasné co znamenají první a druhý sloupeček "dat" pod grafem, neboť chybí popis i legenda. Spojnicový graf nedává v daném kontextu smysl.

V kapitole 1 se autor často snaží definovat již zavedené pojmy sám, podobně tak i zkouší sám sepisovat tvrzení a jejich důkazy, místo aby převzal ty ustálené (např. Def 1.1.6, Def. 1.2.5, ...). Definice 1.1.6 je podivná, standardní literatura definuje substituci dvou stromů, ne stromu a seznamu substitucí za jednotlivé substituční symboly. Podle této definice mi pak nedává smysl definice operace $L_1 \cdot L_2$. Definice 1.2.5 má být zapsána jako algoritmus.

Definice 5.2.1 a Lemma 5.2.1 neřeší ohodnocený graf, kterým FTA je.

Rešerše se omezuje na články [10] a [9]. Místo [9] by bylo lepší citovat Laugerotte, Ouali-Sebti, Ziadi: From Regular Tree Expression to Position Tree Automaton. (LATA 2013). Diplomová práce [9], ani z ní vycházející článek (necitováno), nepřevádí na FTA, nýbrž na PDA, nicméně metoda lze lehce upravit, aby dávala FTA. V rešerši se autor vůbec nezabývá ostatními převody z RTE na FTA a zpět (např. Kuske and Meinecke 2008, Belabbaci et al. 2018). Reference [2] odkazuje na celý sborník, místo na konkrétní článek. Odkazy [3] a [4] vedou na stejnou práci.

K novým algoritmům ze sekcí 4.1 a 4.2 bohužel neexistují žádné příklady, které by čtenáři usnadnily pochopení algoritmů.

V sekci 5 se vyskytuje věta "Algorithm 4.2.2 is implemented...". Algoritmus 4.2.2 neexistuje, pravděpodobně je myšlen algoritmus ze sekce 4.2.2. Nicméně naimplementovaný jsem našel i algoritmus ze sekce 4.1, o tom zde není zmínka

Na druhou stranu, autorovy výsledky nejsou vůbec špatné, pokulhává ale jejich prezentace.

Další poznámky:

- p1: Glushkow -> Glushkov
- Někde mezi definicemi 1.1.6 a 1.1.8 chybí definice zřetězení nad stromovými jazyky
- Def. 1.1.7: Místo E má být L
- Def. 1.3.1.: Místo \square má být \square_x
- Theorem 1.3.3: [9] není validní zdroj pro tuto větu, viz poznámka k rešerši
- Sekce 2: "I can mention the source [3]" -> lepší asi "I can mention the work of Drewes [3]", to samé i dále
- Sekce 4.1, Alg. 1: Kód obsahuje podivné znaky namísto asi zamýšlených $\langle a \rangle$
- Sekce 4.2.2: finite states -> final states
- Sekce 4.2.2, p.36: Autor se odkazuje na řádky neznámého algoritmu (očividně Algorithm 6)
- Sekce 4.2.2, Alg. 6: Používá se nedefinovaná funkce FinalToRTE
- Sekce 5.3.2, 5.3.3: 2st -> 2nd, 3st -> 3rd
- Sekce 5.4: FTE -> RTEs

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů
(známka A až F):

3. Nepísemná část, přílohy

89 (B)

Popis kritéria:

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

Komentář:

Jako přílohu hodnotím implementované algoritmy v ALT.

Identifikátory by si zasloužily lepší jména (kids, leadingToMe), překlepy se nevyhly ani jim (incomming).

Kód bych radši viděl rozdělený do menších logických celků s komentáři. Některé metody jsou celkem dlouhé, pravděpodobně by pomohlo rozdělení do logických bloků.

Kód je v pořádku a otestovaný.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

4. Hodnocení výsledků, jejich využitelnost

90 (A)

Popis kritéria:

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.

Komentář:

Hlavními výsledky práce jsou nové poznatky o převodu FTA na RTE a jejich implementace. Bude potřeba dále vyzkoumat, možná i navazujícími diplomovými pracemi, zda nejdou obě autorem navrhované metody spojit, a jak tyto poznatky souvisí s převodem pomocí regulárních stromových rovnic. Zajímavé jsou i naměřené rozdíly ve velikosti RTE mezi implementovanými metodami.

Práce jistě poslouží dalšímu výzkumu, některé myšlenky budou jistě dále rozpracovány.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).

Otázky:

- Máte nějaké nápady, jak zmenšit složitost druhého algoritmu (sekce 4.2.2)?

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

6. Celkové hodnocení

56 (E)

Popis kritéria:

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.

Text hodnocení:

Dosažené výsledky jsou dobré a autor projevilschopnost se zorientovat v tématu a přijít s novým algoritmem. Nicméně kvality textové části práce se velmi projevily v mém hodnocení.

Práci doporučuji k obhajobě a hodnotím 56 body, tedy známkou E (dostatečně).

Podpis oponenta práce: