

СТАНОВИЩЕ

във връзка с процедура за защита на дисертационен труд
за придобиване на образователната и научна степен „доктор“

по: област на висшето образование 4. *Природни науки, математика и информатика*, професионално направление 4.6. *Информатика и компютърни науки*, докторска програма *Информатика*

Автор: **Николай Георгиев Ханджийски**, редовен докторант към катедра „Компютърна информатика“ във Факултета по математика и информатика при ПУ „П. Хилендарски“

Тема: **Итеративен парсиращ алгоритъм**

с приложение в профилирането на парсери

Научен ръководител: **проф. д-р Елена Петрова Сомова**

Автор на становището: Галина Мирчева Илиева, професор в професионално направление: 3.8. Икономика, доктор по Информатика, Пловдивски университет „Паисий Хилендарски“

1. Общо представяне на процедурата и докторанта

Настоящото становище е изготвено на основание Заповед № РД–21–654/21.03.2024 г. на Ректора на Пловдивски университет “Паисий Хилендарски“ (ПУ) – проф. д-р Румен Младенов въз основа на решение (Протокол № 5/20.03.2024 г.) на ФС на Факултета по математика и информатика (ФМИ) и доклад на проф. д-р Ангел Атанасов Голев, декан на ФМИ, в съответствие с Чл. 4. на ЗРАСРБ, Чл. 2. (2), Чл. 30. (3) на Правилника за приложение на ЗРАСРБ и Чл. 37. (1) на Правилника за развитие на академичния състав на Пловдивски университет „Паисий Хилендарски“ (ПРАСПУ). С цитираната заповед съм назначена за член на научното жури за осигуряване на процедура за защита на дисертационен труд на тема „Итеративен парсиращ алгоритъм с приложение в профилирането на парсери“ за придобиване на образователната и научна степен „доктор“ в Област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки, по докторска програма Информатика. Научен ръководител е проф. д-р Елена Петрова Сомова.

Николай Георгиев Ханджийски се дипломира като Магистър (Софтуерни технологии) във ФМИ на ПУ „П. Хилендарски“. Владее английски и немски език.

2. Актуалност на тематиката и целесъобразност на поставените цели и задачи

Усъвършенстването на алгоритмите за парсиране е актуален проблем на съвременната информатика поради няколко причини. Първо, те подобряват процеса на разработка на софтуер чрез предоставяне на разширени възможности за синтактичен анализ. Вградени в парсери, те позволяват на разработчиците да откриват грешки по-ефективно и предлагат функции за автоматично допълване на код (autocomplete) в интегрираните среди за

разработка (Integrated Development Environments – IDEs). Тези промени водят до подобро качество на кода, по-бърз цикъл на разработка и повишена продуктивност на разработчиците. Второ, тези алгоритми играят съществена роля в компютърната лингвистика, като позволяват точен синтактичен анализ и коректно семантично разбиране на езика. Те решават различни задачи, например машинен превод и извличане на информация, като допринасят за разработването на сложни системи за езикова обработка в различни области.

Дисертационният труд е посветен на изследването на теоретичната основа и практическото приложение на парсери за конкретен тип многозначни безконтекстни граматика. Създадени са: 1) специална теоретична рамка; 2) парсираща машина; 3) парсиращ алгоритъм и 4) профилатор. Последният улеснява верификацията на авторските идеи като генерира контекстно-свободни граматика и входни данни за експерименти с тях при автоматизирано отчитане на използваните изчислителни ресурси при парсирането. Докторското изследване обхваща шест задачи.

3. Познаване на проблема

Николай Ханджийски проучва са множество литературни източници в изследваната проблематика. Той познава добре основните аспекти на проектирането на транслатори и по-специално алгоритмите за разпознаване и парсиране на езици, дефинирани от безконтекстни граматика.

4. Методика на изследването

За постигане на целите в дисертационния труд докторантът следва следната схема: 1) проучване и анализиране на специализирана литература, свързана с предмета на изследване; 2) разработване на парсираща машина за изследване на възможностите на различни парсери; 3) представяне на алгоритъма за тунелно парсиране и обектите, необходими за работата му; и 4) верифициране на предложената методика чрез серия от експеримента, осъществени чрез специално създаден за целта инструмент – профилатор.

Методиката, предложена от докторант Николай Ханджийски, е подходяща и позволява да се решат поставените задачи и да се постигнат набеязаните цели.

5. Характеристика и оценка на дисертационния труд и приносите

Дисертационният труд съдържа общо 181 страници и е структуриран в увод, изложение в четири глави, заключение и списък на използваната литература. Основният текст е 163 страници, като допълнително има представени и две приложения по повод работата на

профилатора, които са в обем от две страници. Използваните източници са общо 191 на брой, от които 187 на английски език, 2 на руски език и 2 на български език.

В първа глава са дефинирани основни термини от теорията на формалните езици и абстрактните машини, свързани с процеса на транслиране и обработка на програмен код. Представени са особеностите и приложенията на крайните автомати и регулярните изрази, класифицирани са видовете граматика, описани са съставните части на трансляторите и е направен обзор на съществуващите алгоритми за парсиране на безконтекстни граматика. Втора глава съдържа подробно описание на предложената парсираща машина. Тя е от хибриден тип, т.е. позволява парсиране без или с лексикален анализ. Всеки елемент на парсиращата машина е описан подробно на основата на неговите функционални възможности. В Трета глава е представен псевдокодът на нов парсиращ алгоритъм за неляво рекурсивни напреднали (усъвършенствани) граматика (non-left recursive advanced grammars) – „тунелно парсиране“ (Tunnel Parsing – TP). Алгоритъмът е демонстриран на базата на конкретен пример при зададен низ за входни данни. Четвърта глава описва софтуерен инструмент, наречен профилатор, за генериране на контекстно-свободни граматика и входни данни за парсиране. Допълнително, този инструмент подпомага отчитането на експерименти с парсиращи машини. В главата се представя и новосъздаденият език за създаване на шаблонни граматика. Анализирани са резултатите, получени от четирите експеримента, проведени с помощта на профилатора. В Заключението са обобщени и систематизирани получените резултати, като са посочени основните научни, научно-приложни и приложни приноси на дисертационния труд. Формулирани са перспективи за бъдещо развитие на дисертационната тематика.

Получените резултати потвърждават хипотезата на дисертационната работа. Те са значими и представляват принос в разработването на ефективни парсери, базирани на някои видове напреднали граматика, с линейна сложност по време.

Основните приноси на докторанта са научни, научно-приложни и приложни. Научните приноси на дисертационното изследване са следните: Разработен е концептуален модел на парсираща машина с разширени възможности спрямо съществуващите аналози. Дефинирани са напреднали граматика с фраза символи като съставени от правила и напреднали символи чрез нотация, подобна на подсилената (augmented) Бекус-Наурова форма. Предложен е модел на фразова машина, който ускорява анализа като предварително категоризира фрази от парсер граматиката. Научно-приложните приноси са: описана е функционалността на парсиращата машина; създаден е алгоритъм „тунелно парсиране“, парсиращ на базата на

неляво рекурсивни напреднали граматика при линейна сложност по време; създаден е език за мета-програмиране на граматика; проектиран и реализиран е профилатор на парсер генератори, позволяващ извършването на тестове с парсиращи машини, създадени от различни парсер генератори и отчитане на използваните изчислителни ресурси. Приложните приноси на дисертационното изследване са: 1) разработен е прототип на софтуерен инструмент – профилатор на парсер генератори за експерименти както с директно въведена граматика, така и със съвкупност от граматика, програмно генерирани чрез самия инструмент, с вграден модул за визуализация на резултатите и 2) с помощта на профилатора са проведени експерименти за оценяване на ефективността и сложността на парсери.

6. Преценка на публикациите по дисертационния труд и личния принос на докторанта

Във връзка с дисертационния труд докторантът е представил 6 публикации. Шестте публикации са на английски език и са в съавторство, като Николай Ханджийски е първи автор. Четири от публикациите са статии в списания, а останалите две са публикувани в сборници на научни конференции. Всички публикации са индексирани във световните наукометрични бази от данни – Web of Science и/или Scopus. Впечатляващ е и фактът, че две от статиите са публикувани в списание с импакт фактор.

Убедена съм, че предложеният дисертационен труд е лично дело на автора.

Не е установено доказано по законоустановения ред плагиатство в представените по конкурса научни трудове.

7. Оценка на автореферата

Представеният автореферат е изготвен съгласно изискванията на закона и съответните правилници, като представя постигнатите основни резултати и приноси.

8. Препоръки за бъдещо използване на дисертационните приноси и резултати

Бих препоръчала на Николай Ханджийски да реализира декларираните от него перспективи за бъдещо развитие на изследванията.

ЗАКЛЮЧЕНИЕ

От гореизложеното се вижда, че Николай Георгиев Ханджиев е висококвалифициран експерт в областта на информатиката с натрупан опит и доказани постижения в тази област. Той притежава знания и умения за самостоятелно провеждане на научни изследвания. Това ми дава основание да заключа, че освен изпълнените минимални национални изисквания,

дисертационният труд, авторефератът и представената научна продукция удовлетворяват изискванията на ЗРАСРБ, на Правилника за неговото прилагане, както и на Правилника на ПУ „Паисий Хилендарски“ и специфичните изисквания на Факултета по математика и информатика при ПУ за развитие на академичния състав, предявявани към кандидати за придобиване на образователната и научна степен „доктор“.

Поради това **давам своята положителна оценка** за проведеното изследване, представено от рецензираните по-горе дисертационен труд, автореферат, постигнати резултати и приноси, и предлагам на почитаемото научно жури да присъди образователната и научна степен „доктор“ на Николай Георгиев Ханджийски в Област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки, докторска програма Информатика.

09.05.2024 г.

Изготвил становището:

гр. Пловдив

(проф. д-р Галина Илиева)