

## СТАНОВИЩЕ

от доц. д-р Юлиана Пенева Пашкова

преподавател в Нов български университет, деп. “Информатика“

на дисертационен труд за присъждане 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.

докторска програма *Информатика*

**Автор:** *Николай Георгиев Ханджийски*

**Тема:** *„Итеративен парсиращ алгоритъм с приложение в профилирането на парсери“*

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

*Пловдивски университет „П. Хилендарски“*

### 1. Общо представяне на процедурата и докторанта

Със заповед № РД-21-654 от 21.03.2024 г. на Ректора на Пловдивския университет „Паисий Хилендарски“ (ПУ) съм определена 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, докторска програма Информатика. Автор на дисертационния труд е Николай Георгиев Ханджийски – докторант в редовна форма на обучение към катедра Компютърна информатика с научен ръководител проф. д-р Елена Петрова Сомова от ПУ „Паисий Хилендарски“ - ФМИ.

Представеният комплект материали на хартиен носител е в съответствие с Чл.36 (1) от Правилника за развитие на академичния състав на ПУ, и съдържа:

- молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд;
- автобиография в европейски формат;
- протокол от катедрения съвет, свързан с докладване на готовността за откриване на процедурата и с предварително обсъждане на дисертационния труд;
- автореферат – на български и на английски език;
- декларация за оригиналност и достоверност на приложените документи;
- справка за спазване на минималните национални изисквания;
- списък на научните публикации по темата на дисертацията;
- дисертационен труд;
- копия на научните публикации;

Приложени са 6 публикации на докторанта.

Докторант Николай Ханджийски е магистър по информатика на ФМИ. По време на обучението си е бил студент по програма Еразъм в университет Кобленц-Ландау, Кобленц, Германия и студент по програма Цепус в университет Йоханес Кеплер, Линц, Австрия. Понастоящем работи като софтуерен инженер в българска фирма, разработваща софтуер за управление на търговски обекти и като управител на немска фирма със същия предмет на дейност. Има педагогически опит като хоноруван преподавател по информатични дисциплини в ФМИ на ПУ както и участие в научни изследвания в университет Кобленц-Ландау, Кобленц, Германия.

### 2. Актуалност на тематиката

Дисертационният труд е посветен на автоматизирането на процеса парсинг, свързан с извличането на структурирана информация от неструктурирани данни. Парсирането на конкретни данни означава разпознаване на тяхната принадлежност към определен език и извеждането на конкретна структурна информация за тях. Парсиращите алгоритми изграждат структура от данни, която може да бъде използвана от компилатор, интерпретатор или преводач за създаване на изпълнима програма или библиотека. Те са ключови за анализа и обработката на езикови структури, особено в контекста на увеличаващото се количество данни, генерирани от различни приложения и платформи. Предмет на настоящото научно

изследване е разработването на средства (парсиращи машини и алгоритми), които са подходящи за линейно транслиране на данни на базата дадена граматика. Предвид приложимостта на тези средства за развитието на технологиите и интересът на водещи компании към тях, намирам тематика на дисертационния труд за актуална.

Цел на дисертационното изследване е проектиране на парсираща машина, в която да се използват различни парсиращи алгоритми, за да се добавят нови функционалности. Изследва се съответния клас от граматика, при които предложения от докторанта Тунелен алгоритъм, дава резултати за линейно време. Разработен и тестван експериментално е прототип на профилатор на парсер генератори, представляващ инструмент за измерване на ресурсите, които се използват по време на разпознаване/парсиране на данни (породени от различни безконтекстни граматика) от автоматично генерирани парсери с различни парсер генератори и компилатори. Поставената цел и произтичащите от нея задачи са в съответствие със съвременността на проблема.

### **3. Познаване на проблема**

Прегледът на цитираната литература (общо 191 заглавия, 4 на кирилица и 188 на латиница, от тях 10 са интернет източници) позволява да се твърди, че докторантът е навлязъл достатъчно добре в проблематиката. Списъкът е представителен по брой, по години и по разпределение на авторите.

Изготвеният от докторанта обзор на формалните средства, свързани с математическите основи на транслаторите и езиците за програмиране показва задълбочено проучване на състоянието на изследванията в разглежданата област. Представени са основните понятия като комбинаторни схеми, регулярни изрази, безконтекстни граматика, крайни и стекови автомати. Проучени са транслаторите като средство за преобразуване на програми от един формален език на друг и свързаните с тях лексикален и синтактичен анализ. Изследвани са възможностите на различни компилатори, лексер и парсер генератори. Авторът е проучил задълбочено проблематиката на дисертационното изследване.

### **4. Методика на изследването**

Методиката, приложена от докторант Николай Ханджийски произтича от поставените цели и обособените изследователски задачи. Налице са методическите компоненти на дисертационна разработка с приложна насоченост: анализ на състоянието, теоретичен модел, софтуерна реализация, експеримент, доказващ използваемостта на разработените инструменти.

След проведеното проучване в областта и дефиниране на проблема е създадена подходящо обоснована архитектура на парсираща машина и е дефиниран нов вид безконтекстни граматика. Разработен е съответен алгоритъм за парсиране, вграден в парсиращата машина. Предложен е прототип на инструмент за измерване и сравняване на използваните ресурси от различни парсери и са осъществени серия от експерименти с него, с което се доказва практическата приложимост и използваемост на разработката. Може да се твърди, че докторантът е използвал методически похвати, присъщи на коректно научно изследване.

### **5. Характеристика и оценка на дисертационния труд и приносите**

Дисертацията с общ обем от 181 страници съдържа следните компоненти: списък на таблиците, списък на фигурите, таблица с използвани термини от литературата, таблица с въведени термини в дисертацията, списък на използваните съкращения, увод, четири глави и заключение, приноси, перспективи, списък на авторските публикации по темата, списък на забелязани цитирания, участия в проекти, приложения, списък на използваната литература и декларация за оригиналност.

Работата е илюстрирана с подходящо избрани фигури (общо 56) и таблици (общо 8), и като обем е напълно в нормите. Формулираните шест задачи произтичат от целта на дисертационното изследване: да се разработят средства, които са подходящи за линейно транслиране на данни на базата на някои многозначни безконтекстни граматика. Приносите са научни, научно-приложни и приложни, като обхващат: разработване на концептуален модел на парсираща машина с нови функционалности спрямо използваните досега машини; дефиниране на специфични граматика с фраза символи; създаване на алгоритъм тунелно парсиране с линейно време за парсиране за определени граматика; проектиране на език за

метапрограмиране на граматика; реализиране на парсер генератор, позволяващ извършването на тестове (измерване на използваните ресурси по време на разпознаване и парсиране) с парсиращи машини, създадени от различни парсер генератори и компилатори.

Перспективите за развитие на тематиката обхващат: разширяване на списъка от команди, които се изпращат от парсер модула и се приемат от следващите модули, за да се генерират различни от синтактичните дървета синтактични структури; надграждане на предложението Тунелен алгоритъм с цел парсиране на базата на граматика с лява рекурсия; добавяне на нови функционалности към профилатора (поддържане на допълнителни парсер генератори и компилатори); измерване на различни характеристики на процеса „парсиране“.

#### **6. Преценка на публикациите и личния принос на докторанта**

Докторантът е представил списък от шест публикации по дисертационното изследване: 4 - в специализирани списания, 1 - в трудовете на международна конференция и 1 – в трудовете на международен форум. Публикациите са индексирани: 4 - в Web of Science и 5 - в Scopus, общо 5 са в издания с SJR, като в тях са отразени основните резултати, получени в дисертационния труд. Всички публикации са в съавторство с научния ръководител. Забелязан е цитат на 1 от публикациите по темата в научно изследване, което е индексирано в световните бази от данни.

Резултатите са представени в достатъчна степен пред специализирана научна аудитория. Като брой публикации са достатъчни. Изпълнени са специфичните изисквания на ФМИ. Тематиката на публикациите подчертава личното участие на докторанта.

#### **7. Автореферат**

Авторефератът е съставен според изискванията и отразява достатъчно пълно всички аспекти на дисертационното изследване. Обобщени са основните постигнати резултати и приносите на автора.

#### **8. Препоръки за бъдещо използване на дисертационните приноси и резултати**

Парсиращите алгоритми са базово средство за обработка на данни. Дисертационните приноси и резултати могат да намерят приложение при всякакви процеси за анализ и структуриране на входни данни.

#### **ЗАКЛЮЧЕНИЕ**

Дисертационният труд *съдържа научни, научно-приложни и приложни резултати, които представляват оригинален принос в науката и отговарят на всички* изисквания на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за прилагане на ЗРАСРБ и съответния Правилник на ПУ „Паисий Хилендарски“. Представените материали и дисертационни резултати **напълно** съответстват на специфичните изисквания на Факултета по математика и информатика, приети във връзка с Правилника на ПУ за приложение на ЗРАСРБ.

Дисертационният труд показва, че докторант Николай Георгиев Ханджийски **притежава** задълбочени теоретични знания и професионални умения по научна специалност информатика като **демонстрира** качества и умения за самостоятелно провеждане на научно изследване. Резултатите и приносите са оригинални и не са заимствани от чужди изследвания и публикации.

Поради гореизложеното, убедено 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“.

01.05.2024 г.

Изготвил становището: .....  
доц. д-р Юлиана Пенева