**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра технологий программирования**

Аннотация к магистерской диссертации

**Анализ исходного кода на предмет заимствований в
рамках процесса автоматизированного контроля знаний**

Ярошевич Яна Олеговна

Научный руководитель - заведующий кафедрой ТП ФПМИ, доктор
технических наук, профессор Курбацкий Александр Николаевич

2019

РЕФЕРАТ

Магистерская диссертация, 78 с., 17 рис., 14 табл., 36 источников.

Ключевые слова: АЛГОРИТМЫ ПОИСКА ПЛАГИАТА, ИДЕНТИФИКАЦИЯ ЗАИМСТВОВАНИЙ, АЛГОРИТМЫ ПРЕДСТАВЛЕНИЯ ИСХОДНЫХ КОДОВ ПРОГРАММ, ТОКЕНИЗАЦИЯ, НАХОЖДЕНИЕ НАИБОЛЬШЕЙ ОБЩЕЙ ПОДСТРОКИ, КОЭФФИЦИЕНТ ЖАККАРА, МЕТОД ПРОСЕИВАНИЯ, МЕТОД ОТПЕЧАТКОВ, МЕТРИКИ, КОМБИНИРОВАННЫЙ АЛГОРИТМ, АВТОМАТИЗИРОВАННАЯ СИСТЕМА КОНТРОЛЯ ЗНАНИЙ, ПРОВЕРКА КОДА НА ЧИСТОТУ.

Объект исследования – алгоритмы идентификации заимствований, системы контроля знаний, подходы к проверке кода на «чистоту».

Цель работы – исследовать существующие алгоритмы идентификации заимствований, выполнить их сравнение и анализ, на базе исследований выработать комбинированный алгоритм. Реализовать комбинированный алгоритм, протестировать его на студенческих работах. Спроектировать и разработать автоматизированную систему контроля знаний, включающую практическую и теоретическую части. Рассмотреть подходы к проверке исходного кода на «чистоту».

Результатом является автоматизированная система контроля знаний, включающая практическую и теоретическую части, и реализованный комбинированный алгоритм идентификации заимствований.

Область применения – программирование, учебный процесс.

ABSTRACT

The master's thesis, 78 pages, 17 figures., 14 tables, 36 literature references.

Keywords: PLAGIARISM DETECTION ALGORITHMS, BORROWINGS IDENTIFICATION, SOURCE CODES REPRESENTATIONS, TOKENIZATION, SEARCH OF LARGEST COMMON STRING, JACCARD INDEX, SCREENING METHOD, DOCUMENT FINGERPRINTING METHOD, METRICS, COMBINED ALGORITHM, AUTOMATED KNOWLEDGE CONTROL SYSTEM, VERIFICATION CODE ON CLEANLINESS.

Object of research – plagate search algorithms, automated knowledge control systems, approaches for checking the source code on "cleanliness".

Purpose – explore existing algorithms for identifying borrowings, perform their comparison and analysis, and develop a combined algorithm based on research. Implement the combined algorithm, and then test it on student works. Design and develop an automated knowledge control system that includes practical and theoretical parts. Approaches for checking the source code on "cleanliness" are considered.

Result is an automated knowledge control system that includes practical and theoretical parts, and the implemented combined algorithm for identifying borrowings.

Area of application - programming, educational process.