

Голові разової
спеціалізованої вченої ради
Харківського національного
університету імені В. Н. Каразіна
професору Олені ТОЛСТОЛУЗЬКІЙ
майдан Свободи 4, м. Харків, 61022

Рецензія

офіційного рецензента, в.о. завідувача кафедри теоретичної та прикладної інформатики факультету математики і інформатики, Харківського національного університету імені В. Н. Каразіна, кандидата технічних наук, доцента Меньяйлова Євгена Сергійовича на дисертаційну роботу Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», подану на здобуття наукового ступеня доктора філософії з галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки.

1. Обґрунтування вибору теми дослідження.

Вибір теми дослідження, зосередженої на функціональному програмуванні та оптимізації стратегій редукції в безтиповому лямбда-численні, визначається як теоретичною, так і практичною значущістю в сучасній розробці програмного забезпечення. Функціональне програмування використовується все більш широко через його внесок у надійність програмного забезпечення, перевірку, валідацію та масштабованість. Дослідження аналізує методи оцінки складності скорочення лямбда-термів, враховуючи як час для окремих кроків, так і загальну кількість кроків.

Дослідження дає змогу глибше зрозуміти процес скорочення в лямбда-численні, основоположному елементі функціонального програмування. Воно використовує методи машинного навчання, такі як лінійна регресія та штучні нейронні мережі (ШНМ), для прогнозування часу редукції на основі характеристик терму.

Оптимізуючи стратегії редукції, дослідження спрямоване на підвищення обчислювальної продуктивності функціональних програм. Використання сучасних методів машинного навчання для аналізу факторів, що впливають на час редукції, представляє новий підхід у цій галузі.

Дослідження розглядає критичні проблеми в рамках парадигми та вводить нові підходи, які можуть підвищити ефективність і продуктивність функціональних мов програмування та компіляторів. Підсумовуючи, тему дослідження було обрано через її потенціал впливу на функціональне програмування. Все вищесказане підтверджує про актуальність теми дисертаційного дослідження.

2. Оцінка змісту дисертації, її завершеності в цілому і оформлення.

Дисертація складається з анотацій, вступу, чотирьох розділів, висновків, списку використаних джерел та одного додатку. Загальний обсяг роботи складає 174 сторінки друкованого тексту: 143 сторінки основного тексту, 30 рисунків, 16 таблиць, списку використаних джерел із 105 найменувань.

У **вступі** обґрунтовано вибір теми дослідження, визначено мету і основні задачі, об'єкт і предмет дисертаційного дослідження, визначено наукову новизну роботи і практичне значення отриманих результатів, описано методи, які були використані у дослідженнях і особистий внесок здобувача.

У **першому розділі** представлено теоретичну частину дослідження, де детально описано мови функціонального програмування та лямбда-числення, а також зв'язок між ними. Представлено бібліотеку Lambda Calculus Environment, яка використовується як основне середовище для проведення експериментів..

У **другому розділі** описано змішані та рандомізовані стратегії редукції, представлено результати експериментів та порівняння із детермінованими стратегіями. Крім того, пояснюється метод прогнозування часу необхідного для одного кроку редукції лямбда-терму. У розділі розглядається використання даних про стан терму до і після процесу редукції для оцінки часу. Показано, що точна оцінка обчислювальних витрат може стати основою жадібної стратегії з мінімізацією витраченого часу на процес редукції.

У **третьому розділі** було розроблено метод прогнозування кількості кроків редукції лямбда-термів за заданою стратегією із застосуванням методів глибинного навчання. Додатково було розглянуто можливість використання вбудовувань для репрезентації різниці в редукції лямбда-термів при різних стратегіях. Для цього було протестовано чотири LLM-моделі, що генерували вбудовування на основі текстових представлень лямбда-термів.

У **четвертому розділі** представлено результати застосування методів неконтрольованого машинного навчання. Кластеризацію даних було досліджено за допомогою методу DBSCAN і чотирьох моделей вбудовування. Після цього було оцінено інформативність усереднених змінних вбудовувань, введено коефіцієнт перекриття та запропоновано новий підхід до використання LLM безпосередньо в процесі редукції лямбда-терму.

Висновки за результатами виконання дисертаційної роботи підкреслюють наукову новизну та практичну цінність проведених досліджень.

Список використаних джерел свідчить про те, що під час роботи було проаналізовано сучасні результати наукових досліджень.

Дисертація є завершеною науковою працею, а її оформлення відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 12.01.2022 р. № 44), та наказу Міністерства освіти і науки України від 12.01.2017 р. № 40 «Про затвердження Вимог до оформлення дисертації».

3. Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертаційна робота відповідає науковому напрямку кафедри теоретичної та прикладної інформатики факультету математики і інформатики Харківського національного університету ім. В.Н. Каразіна.

Основні результати роботи було реалізовано в Харківському національному університеті імені В. Н. Каразіна у рамках НДР «Моделювання інформаційних процесів у складних і розподілених системах» за 2021 – 2023 рр. (ДР No 0121U109183).

4. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації.

Обґрунтованість наукових положень, висновків і рекомендацій забезпечено коректним використанням архітектур моделей штучного інтелекту, методів імітаційного та математичного моделювання, теорії графів, та методів кластерного аналізу.

Достовірність і обґрунтованість отриманих результатів забезпечується:

- 1) використанням відомих та надійних архітектур нейронних мереж, включаючи LSTM, CNN і Transformer;
- 2) використанням спеціально розробленого та перевіреного програмного забезпечення для експериментальної оцінки;
- 3) проведенням експериментів з даними, отриманими за допомогою методу, який гарантує статистичну значущість;
- 4) об'єктивним порівнянням шляхом використання стандартизованих наборів даних у різних експериментах;
- 5) позитивним сприйняттям результатів науковою спільнотою, відображеним в публікаціях та виступах на міжнародних наукових конференціях.

5. Основні наукові результати, одержані автором, та їх новизна.

Основні наукові результати та висновки дисертації пройшли апробацію під час міжнародних наукових конференцій та знайшли відображення в публікаціях у фахових та міжнародних наукових виданнях. За результатами дисертаційного дослідження опубліковано 7 наукових праць, серед яких: 2 публікації у міжнародних виданнях, які входять до наукометричних баз Scopus та Web of Science; 2 публікації у науковому виданні, включеному на дату опублікування до переліку наукових фахових видань України; 3 праць – у матеріалах і тезах конференцій;

Дисертант отримав наступні наукові результати:

- 1) *вперше розроблено* параметричну випадкову однокрокову стратегію редукції, що відрізняється від існуючих можливістю налаштування ймовірності вибору редексу для скорочення. Варіація параметрів дозволяє керувати процесом редукції лямбда-термів з метою підвищення його продуктивності для конкретних підмножин термів.
- 2) *удосконалено* метод оцінки складності кроку редукції лямбда терму, що відрізняється від існуючих використанням методів машинного навчання. Це дозволило підвищити точність прогнозування обчислювальної складності кроку редукції довільного лямбда-терму.
- 3) *удосконалено* експериментальне середовище емуляції лямбда-числення завдяки впровадженню процедури багатоетапної верифікації розробленого програмного забезпечення та автоматичної генерації лямбда-термів. Це дозволило забезпечити коректну реалізацію цього програмного забезпечення та створювати набори

даних лямбда-термів для забезпечення теоретично необхідного покриття контрольованими конфігураціями лямбда-термів.

- 4) *дістали подальшого розвитку* методи представлення лямбда-термів у вигляді вбудовувань (embeddings), що на відміну від існуючих представлень забезпечує використання штучних нейронних мереж архітектури типу Transformer. Це дозволило збільшити глибину аналізу коду лямбда-термів та підвищити точність екстракції характеристик, що впливають на процес редукції.
- 5) *дістав подальшого розвитку* комплексний метод оптимізації стратегії редукції, що на відміну від існуючих поєднує моделі штучного інтелекту для оцінки часу (в кількості кроків) редукції з процедурою вибору стратегії нормалізації терму. Цей комплексний підхід є новим і в має потенціал скоротити загальний час нормалізації, що суттєво сприяє оптимізації програм, реалізованих функціональними мовами програмування.

6. Практичне значення одержаних результатів.

Представлені в рамках роботи дослідження, хоча здебільшого теоретичні, мають практичне значення у сфері функціонального програмування та оптимізації компіляторів. Було розроблено параметричну однокрокову випадкову стратегію редукції термів лямбда-числення, яка безпосередньо впливає на підвищення ефективності функціональних програм з точки зору часу виконання. Запропоновано метод прогнозування часу редукції окремих редексів і загальної кількості кроків редукції, що допомагає автоматизувати процес вибору стратегії редукції.

Незважаючи на деякі обмеження в точності підходу, робота створює основу для майбутніх досліджень. Можливість передбачати кількість кроків редукції лямбда-термів за допомогою навчених моделей має критичне значення для автоматизації пошуку більш продуктивних стратегій редукції. Окрім функціонального програмування, дослідження також встановлює зв'язки між декількома сферами, такими як теорія лямбда-числення, машинне навчання та розробка програмного забезпечення.

Результати дисертаційного дослідження Дейнега О. А. були впроваджені у навчальний процес Харківського національного університету імені В. Н. Каразіна.

7. Дотримання академічної доброчесності.

За результатами аналізу дисертаційної роботи та публікацій автора порушення академічної доброчесності не виявлено. Елементи фальсифікації чи фабрикації тексту в роботі відсутні.

8. Дискусійні положення та зауваження до змісту дисертації.

Відзначаючи новизну, важливість та повноту одержаних наукових та практичних результатів, слід відзначити й наступні недоліки:

1. У четвертому розділі за допомогою методів PCA та t-SNE представлено візуалізацію прихованого простору датасету лямбда-термів. У наведеній візуалізації можна чітко виділити деякі пусті області в центрі, що свідчить про

можливі прогалини в процесі генерації лямбда термів. Ця особливість графіку є дуже цікавою і, можливо, варта спеціального аналізу.

2. У першому розділі наведено опис бібліотеки Lambda Calculus Environment. Варто було б забезпечити більш детальну специфікацію діаграми класів та, можливо, також подати більш детальний опис основних етапів методу верифікації цього програмного каркасу.
3. В другому розділі представлено метод прогнозування часу виконання одного кроку редукції, для чого використано 3 типи датасетів, та виділено два підходи - один з яких має більш теоретичний характер, а другий - більш прикладний. Можливо, варто було б надати більш детальну аргументацію щодо можливостей та особливостей використання запропонованих методів.
4. В цьому розділі також зазначений вище метод прогнозування часу виконання одного кроку редукції реалізується на платформі операційної системи Unix, а тому варто було б мотивувати такий вибір.
5. У другому розділі, де представлено значення показників MSE, MAE і MAPE для моделей лінійної регресії та ШНМ для трьох наборів даних під час тренування та тестування (табл. 2.2.) вказано доволі велику кількість показників, можливо, варто було б виділи найбільш вагомні з них, та додати більш широкий текстовий опис.
6. У четвертому розділі представлено порівняння значень усереднених вбудовувань із позначенням пріоритету стратегії за алгоритмами стиснення (рис. 4.1.), інформація на графіках є доволі важливою та цікавою з точки зору дослідження, можливо варто було б збільшити масштаб графіків та виділити кожен в окремий рисунок, перенісши їх до додатків.
7. У роботі деякі рисунки та схеми, у вигляді блок-схем, наведено англійською мовою, що ускладнює розуміння. Доцільно було б у пояснювальній частині до рисунків навести означення обраних термінів.

9. Загальні висновки щодо дисертаційної роботи.

Дисертаційна робота Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», представлена на здобуття наукового ступеня доктора філософії галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки є актуальною, завершеною науковою працею, що виконана на належному науково-теоретичному рівні з логічно- та доступно-викладеним матеріалом.

У роботі була розглянута важлива задача дослідження процесу скорочення лямбда-термів і його залежність від структури та розміру терму за допомогою методів машинного навчання. Основним напрямком дослідження є прогнозування часу, необхідного для одного кроку редукції, і прогнозування загальної кількості кроків редукції на основі обраної стратегії. Отримані результати можуть бути застосовані для оптимізації продуктивності компіляторів та інтерпретаторів функціональних мов програмування.

Здобувач Дейнега Олександр Андрійович заслуговує на присудження йому наукового ступеня доктора філософії з галузі знань 12 – «Інформаційні технології» за спеціальністю 122 – «Комп'ютерні науки».

Офіційний рецензент,
кандидат технічних наук, доцент,
в.о. завідувача кафедри теоретичної
та прикладної інформатики
факультету математики і інформатики,
Харківського національного університету
імені В. Н. Каразіна.

Євген МЕНЯЙЛОВ

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ
створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 20:57:05 11.11.2024

Назва файлу з підписом: Рецензія_Меняйлов.pdf.p7s
Розмір файлу з підписом: 222.7 КБ

Перевірені файли:
Назва файлу без підпису: Рецензія_Меняйлов.pdf
Розмір файлу без підпису: 205.5 КБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: МЕНЯЙЛОВ ЄВГЕН СЕРГІЙОВИЧ
П.І.Б.: МЕНЯЙЛОВ ЄВГЕН СЕРГІЙОВИЧ
Країна: Україна
РНОКПП: 3338507797
Організація (установа): ФІЗИЧНА ОСОБА
Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 20:57:05 11.11.2024
Сертифікат виданий: КНЕДП АЦСК АТ КБ "ПРИВАТБАНК"
Серійний номер: 5E984D526F82F38F04000000A6B24201C617C104
Алгоритм підпису: ДСТУ 4145
Тип підпису: Удосконалений
Тип контейнера: Підпис та дані в одному файлі (CAAdES enveloped)
Формат підпису: З повними даними ЦСК для перевірки (CAAdES-X Long)
Сертифікат: Кваліфікований

Версія від: 2024.10.24 15:00

Голові разової
спеціалізованої вченої ради
Харківського національного
університету імені В. Н. Каразіна
професору Олені ТОЛСТОЛУЗЬКІЙ
майдан Свободи 4, м. Харків, 61022

Рецензія

офіційного рецензента, кандидата технічних наук, виконуючого обов'язки директора навчально-наукового Інституту комп'ютерних наук та штучного інтелекту Харківського національного університету імені В. Н. Каразіна Узлова Дмитра Юрійовича на дисертаційну роботу Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», подану на здобуття наукового ступеня доктора філософії з галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки.

1. Обґрунтування вибору теми дослідження.

Дослідження спрямоване на вивчення методів оптимізації стратегій редукції у безтиповому лямбда-численні та їх застосування у функціональному програмуванні для підвищення ефективності виконання програм. Вибір теми обумовлений її теоретичною значущістю та впливом на сучасну індустрію розробки програмного забезпечення. Функціональне програмування, базуючись на математичних засадах та лямбда-численні, дозволяє створювати більш передбачувані програми, за рахунок їх більшої придатності до аналізу.

Незважаючи на зростаючу популярність та переваги, функціональне програмування стикається з викликами, пов'язаними з низькою продуктивністю та неефективним використанням пам'яті. Ці проблеми обмежують його застосування в продуктивних середовищах, де значення має швидкість виконання та ефективне використання ресурсів. Дослідження зосереджене на розробці методів підвищення продуктивності стратегій редукції, які спрямовані на подолання цих недоліків функціональних програм.

Лямбда-числення, як формальна система визначення функцій, є основою багатьох функціональних мов програмування. Дослідження спрямоване на глибше розуміння процесів редукції в лямбда-численні, зокрема, на аналіз складності кроків редукції з точки зору часу виконання та кількості необхідних операцій. Такий підхід дозволяє знайти нові методи оптимізації, які можуть бути інтегровані в компілятори та інтерпретатори, підвищуючи загальну ефективність функціональних програм.

У рамках дослідження використовуються сучасні методи машинного навчання

для прогнозування часу редукції лямбда-термів на основі їх характеристик. Це надає шляхи для автоматизації компіляторів та інтерпретаторів функціональних мов. Використання таких методів, як лінійна регресія та регресійні моделі штучних нейронних мереж (ШНМ), дозволяє підвищити продуктивність програм та їхню ресурсоефективність.

Дослідження надає нові підходи та інструменти для оптимізації стратегій редукції, що робить вагомий внесок у розвиток сучасних мов програмування.

Тема дослідження є актуальною як з теоретичної точки зору, так і з практичної, у галузі функціонального програмування.

2. Оцінка змісту дисертації, її завершеності в цілому і оформлення.

Дисертація складається з анотацій, вступу, чотирьох розділів, висновків, списку використаних джерел та одного додатку. Загальний обсяг роботи складає 174 сторінки друкованого тексту: 143 сторінки основного тексту, 30 рисунків, 16 таблиць, списку використаних джерел із 105 найменувань.

У **вступі** обґрунтовано вибір теми дослідження, визначено мету і основні задачі, об'єкт і предмет дисертаційного дослідження, визначено наукову новизну роботи і практичне значення отриманих результатів, описано методи, які були використані у дослідженнях і особистий внесок здобувача.

У **першому розділі** представлено теоретичну частину дослідження, описані функціональні мови програмування та лямбда-числення, зв'язок між ними. Описано бібліотеку Lambda Calculus Environment, що виконує роль основного оточення для проведення експериментів.

У **другому розділі** представлені змішані та рандомізовані стратегії скорочення з описом експериментальних результатів і порівнянням із детермінованими стратегіями. Крім того, пояснюється метод прогнозування часу, необхідного для одного кроку скорочення лямбда-терму.

У **третьому розділі** було досліджено потенціал для оцінки кількості кроків редукції лямбда-терму на основі заданої стратегії з використанням методів глибокого навчання. Крім того, було досліджено доцільність використання вбудовувань для аналізу відмінностей у процесі редукції лямбда-термів за допомогою різних стратегій. Щоб досягти цього, було протестовано чотири моделі LLM, які генерують вбудовування з текстових представлень лямбда-термів.

У **четвертому розділі** дисертаційної роботи наведено результати застосування неінформованих методів машинного навчання. Досліджено формування кластерів даних за допомогою методу DBSCAN та чотирьох моделей вбудовувань. Після цього проведено оцінку інформативності змінних усереднених вбудовувань, впроваджено коефіцієнт перекриття та запропоновано підхід для використання LLM безпосередньо в процесі редукції лямбда-термів, описані результати вирішення цієї задачі.

Висновки за результатами виконання дисертаційної роботи підкреслюють наукову новизну та практичну цінність проведених досліджень.

Список використаних джерел свідчить про те, що під час роботи було проаналізовано сучасні результати наукових досліджень.

Дисертація є завершеною науковою працею, а її оформлення відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 12.01.2022 р. № 44), та наказу Міністерства освіти і науки України від 12.01.2017 р. № 40 «Про затвердження Вимог до оформлення дисертації».

3. Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертаційна робота відповідає науковому напрямку кафедри теоретичної та прикладної інформатики факультету математики і інформатики Харківського національного університету ім. В.Н. Каразіна.

Основні результати роботи було реалізовано в Харківському національному університеті імені В. Н. Каразіна у рамках НДР «Моделювання інформаційних процесів у складних і розподілених системах» за 2021 – 2023 рр. (ДР No 0121U109183).

4. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації.

Обґрунтованість наукових положень, висновків і рекомендацій забезпечено коректним використанням широкого спектру архітектур моделей штучного інтелекту, методів імітаційного та математичного моделювання, теорії графів, та методів кластерного аналізу.

Достовірність і обґрунтованість отриманих результатів забезпечується:

- 1) використанням усталених і надійних архітектур штучних нейронних мереж, таких як LSTM, CNN і Transformer;
- 2) впровадженням спеціально розробленого та перевіреного програмного забезпечення для проведення експериментів;
- 3) використанням даних, отриманих за допомогою розробленого методу, який гарантує статистичну значущість;
- 4) використанням стандартизованих наборів даних для порівняльного аналізу, для забезпечення об'єктивності оцінювання;
- 5) позитивним визнанням наукової спільноти, про що свідчать публікації та доповіді на трьох міжнародних наукових конференціях.

5. Основні наукові результати, одержані автором, та їх новизна.

Основні наукові результати та висновки дисертації пройшли апробацію під час міжнародних наукових конференцій та знайшли відображення в публікаціях у фахових та міжнародних наукових виданнях. За результатами дисертаційного дослідження опубліковано 7 наукових праць, серед яких: 2 публікації у міжнародних виданнях, які входять до наукометричних баз Scopus та Web of Science; 2 публікації у науковому виданні, включеному на дату опублікування до переліку наукових фахових видань України; 3 праць - у матеріалах і тезах конференцій;

Дисертант отримав наступні наукові результати:

- 1) *вперше розроблено* параметричну випадкову однокрокову стратегію редукції, що відрізняється від існуючих можливістю налаштування ймовірності вибору редексу для скорочення. Варіація параметрів дозволяє керувати процесом редукції лямбда-термів з метою підвищення його продуктивності для конкретних підмножин термів.
- 2) *удосконалено* метод оцінки складності кроку редукції лямбда терму, що відрізняється від існуючих використанням методів машинного навчання. Це дозволило підвищити точність прогнозування обчислювальної складності кроку редукції довільного лямбда-терму.
- 3) *удосконалено* експериментальне середовище емуляції лямбда-числення завдяки впровадженню процедури багатоетапної верифікації розробленого програмного забезпечення та автоматичної генерації лямбда-термів. Це дозволило забезпечити коректну реалізацію цього програмного забезпечення та створювати набори даних лямбда-термів для забезпечення теоретично необхідного покриття контрольованими конфігураціями лямбда-термів.
- 4) *дістали подальшого розвитку* методи представлення лямбда-термів у вигляді вбудовувань (embeddings), що на відміну від існуючих представлень забезпечує використання штучних нейронних мереж архітектури типу Transformer. Це дозволило збільшити глибину аналізу коду лямбда-термів та підвищити точність екстрації характеристик, що впливають на процес редукції.
- 5) *дістав подальшого розвитку* комплексний метод оптимізації стратегії редукції, що на відміну від існуючих поєднує моделі штучного інтелекту для оцінки часу (в кількості кроків) редукції з процедурою вибору стратегії нормалізації терму. Цей комплексний підхід є новим і в має потенціал скоротити загальний час нормалізації, що суттєво сприяє оптимізації програм, реалізованих функціональними мовами програмування.

6. Практичне значення одержаних результатів.

Дослідження, представлені в роботі, хоча переважно теоретичні, мають широкі практичні наслідки в областях функціонального програмування та оптимізації компіляторів. Розроблено параметричну однокрокову випадкову стратегію редукції лямбда-числення, що дозволяє підвищити ефективність функціональних програм з точки зору часу виконання.

У дослідженні описано метод прогнозування часу скорочення конкретних редексів і загальної кількості кроків редукції, пропонуючи ідеї для оптимізації розподілу обчислювальних ресурсів. Хоча представлена модель має певні обмеження щодо точності, дана робота створює основу для майбутніх досліджень. Здатність передбачити кількість кроків редукції для лямбда-термів за допомогою навчених моделей є основою для автоматизації вибору стратегій редукції. Окрім

функціонального програмування, це дослідження пов'язує різні галузі, зокрема теорію лямбда-числення, машинне навчання та розробку програмного забезпечення.

Результати дисертаційного дослідження Дейнега О. А. були впроваджені у навчальний процес Харківського національного університету імені В. Н. Каразіна.

7. Дотримання академічної доброчесності.

За результатами аналізу дисертаційної роботи та публікацій автора порушення академічної доброчесності не виявлено. Елементи фальсифікації чи фабрикації тексту в роботі відсутні.

8. Дискусійні положення та зауваження до змісту дисертації.

Відзначаючи новизну, важливість та повноту одержаних наукових та практичних результатів, слід відзначити й наступні недоліки:

1. У другому розділі, при описі процесу прогнозування часу одного кроку редукції варто було б додати перелік того, що саме впливає на цей час, як саме розмір терма та його структура пов'язані зі складністю обчислення наступного кроку редукції. Також, можливо, мало би сенс додати більший набір характеристик деревного представлення лямбда-терму.
2. В експериментах використовувалися члени згенерованого набору лямбда-термів, результати аналізу яких розглядалися як свідоцтва певних характеристик функціональних мов програмування. Проте згенерований набір може не відображати всіх існуючих залежностей та нюансів реальних функціональних програм. Мало би сенс спробувати використовувати терми, що представляють реальні функціональні програми і можуть бути отримані перетворенням реального функціонального коду, наприклад з проектів з відкритим кодом, у терми лямбда-числення, замість згенерованих лямбда-термів.
3. У роботі представлено графіки навчання моделей, на яких видно ефект перенавчання, та вказано що був використаний підхід Early Stopping, проте даний підхід не можна назвати ідеальним і він має свої особливості. Можливо варто було б більш детально опрацювати архітектуру та визначити розміри моделей.
4. В багатьох розділах, де представлено графіки використаних архітектур моделей машинного навчання, можливо мало би сенс збільшити масштаб або реструктуризувати їх для покращення візуальної складової.
5. В четвертому розділі, де представлено діаграму IDEF0 можливої імплементації описаного підходу можна було б зробити більш детальне візуальне представлення та додати більш широкий опис самого методу в розрізі використання на практиці.

9. Загальні висновки щодо дисертаційної роботи.

Дисертаційна робота Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», представлена на здобуття наукового ступеня доктора філософії галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки є актуальною, завершеною науковою працею, що виконана на належному науково-теоретичному рівні з логічно- та доступно-викладеним матеріалом.

Дане дисертаційне дослідження вирішує важливе наукове завдання, яке полягає у дослідженні процесу редукції лямбда термів, їх залежностей від структури, з використанням методів машинного навчання. Одними з важливих задач, що були вирішені в дослідженні є прогнозування часу одного кроку редукції та прогнозування кількості кроків редукції за обраною стратегією. Отримані результати можуть бути використані для оптимізації роботи компіляторів та інтерпретаторів функціональних мов програмування.

Здобувач Дейнега Олександр Андрійович заслуговує на присудження йому наукового ступеня доктора філософії з галузі знань 12 – «Інформаційні технології» за спеціальністю 122 – «Комп'ютерні науки».

Офіційний рецензент,
кандидат технічних наук,
виконуючий обов'язки директора
навчально-наукового Інституту
комп'ютерних наук та штучного інтелекту
Харківського національного університету
імені В. Н. Каразіна.

Дмитро УЗЛОВ

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ
створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 17:36:33 12.11.2024

Назва файлу з підписом: Рецензія_Узлов_ЄЦП.docx.p7s
Розмір файлу з підписом: 40.8 КБ

Назва файлу без підпису: Рецензія_Узлов_ЄЦП.docx
Розмір файлу без підпису: 22.7 КБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: Узлов Дмитро Юрійович

П.І.Б.: Узлов Дмитро Юрійович

Країна: Україна

РНОКПП: 2613714758

Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 17:31:58
12.11.2024

Сертифікат виданий: "Дія". Кваліфікований надавач електронних довірчих послуг

Серійний номер: 382367105294AF97040000001EC60C0049EED001

Тип носія особистого ключа: ЗНКІ криптомодуль ІІТ Гряда-301

Серійний номер носія особистого ключа: Не визначено

Алгоритм підпису: ДСТУ 4145

Тип підпису: Кваліфікований

Тип контейнера: Підпис та дані в одному файлі (CAAdES enveloped)

Формат підпису: З повними даними ЦСК для перевірки (CAAdES-X Long)

Сертифікат: Кваліфікований

Версія від: 2024.10.11 13:00

Голові разової
спеціалізованої вченої ради
Харківського національного
університету імені В. Н. Каразіна
професору Олені ТОЛСТОЛУЗЬКІЙ
майдан Свободи 4, м. Харків, 61022

Відгук

офіційного опонента, доктора технічних наук, професора кафедри інтелектуальних комп'ютерних систем Національного технічного університету «Харківський політехнічний інститут», професора **Шаронової Наталії Валеріївни** на дисертаційну роботу **Дейнега Олександра Андрійовича** «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», подану на здобуття наукового ступеня доктора філософії з галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки

1. Обґрунтування вибору теми дослідження.

Вибір теми дослідження зумовлений як теоретичною важливістю, так і практичним значенням функціонального програмування та пов'язаної з ним проблеми оптимізації стратегій редукції в безтипovому лямбда-численні для процесу розробки програмного забезпечення. Функціональне програмування відіграє важливу роль у сучасному програмному забезпеченні, сприяючи підвищенню його надійності, за рахунок специфічних методів верифікації, валідації та масштабованості. Дослідження аналізує методи прогнозування складності редукції лямбда-термів шляхом оцінки часу, витраченого на кожен крок, і загальної кількості кроків. Це дослідження покращує розуміння процесу редукції в лямбда-численні, ключовому аспекті функціонального програмування. Дослідження інтегрує як класичні підходи, як от лінійна регресія та дерева ухвалення рішень, так і передові методи машинного навчання, що включають сучасні архітектури штучних нейронних мереж, для передбачення часу скорочення на основі характеристик лямбда-термів.

Оптимізуючи стратегії редукції, дослідження спрямоване на покращення обчислювальної продуктивності функціональних програм з точки зору часу виконання. Застосування сучасних методів машинного

навчання для аналізу факторів, що впливають на час редуції лямбда-термів, є відносно новим підходом у цій галузі. Оскільки індустрія програмного забезпечення все більш широко використовує парадигми функціонального програмування, це дослідження надає сучасні інструменти та методи для вирішення деяких ключових проблем. Представлені методи пропонують перспективні шляхи для автоматизації вибору стратегій редуції, потенційно сприяючи прогресу у функціональному програмуванні. Таким чином, тема дослідження обрана з огляду на її потенційний вплив на сферу функціонального програмування. Все вищесказане свідчить про актуальність теми дисертаційного дослідження.

2. Оцінка змісту дисертації, її завершеності в цілому і оформлення.

Дисертація складається з анотацій, вступу, чотирьох розділів, висновків, списку використаних джерел та одного додатку. Загальний обсяг роботи складає 174 сторінки друкованого тексту: 143 сторінки основного тексту, 30 рисунків, 16 таблиць, списку використаних джерел із 105 найменувань.

У **вступі** обґрунтовано вибір теми дослідження, визначено мету і основні задачі, об'єкт і предмет дисертаційного дослідження, визначено наукову новизну роботи і практичне значення отриманих результатів, описано методи, використані у роботі, і особистий внесок здобувача.

У **першому розділі** представлена теоретична частина дослідження, з детальним описом мов функціонального програмування та лямбда-числення, а також зв'язок між ними. Пояснюється можливість переходу від роботи з функціональними мовами програмування до лямбда-числення. Також описано бібліотеку середовища лямбда-числення, яка використовується як основне середовище для експериментів.

У **другому розділі** представлено змішані та рандомізовані стратегії редуції, з детальним описанням результатів експериментів, та порівнянням їх із детермінованими підходами. Описані результати показують ефективність даного підходу, та можливість заміни чистих стратегій змішаними, що дозволяє зберегти існуючу продуктивність. Даний розділ описує метод, який використовується для прогнозування часу, необхідного для одного кроку скорочення лямбда-терму.

У **третьому розділі** досліджено можливість оцінки кількості кроків редуції лямбда-терму для конкретної стратегії за допомогою методів

глибокого навчання. Досліджено використання вбудовувань для представлення відмінностей у процесі скорочення лямбда-термів при використанні різних стратегій. Протестовано чотири моделі LLM, які генерували вбудовування з текстових представлень лямбда-термів.

У **четвертому розділі** дисертаційної роботи представлено результати застосування неінформованих методів машинного навчання. Аналіз проводився з використанням методів глибинного навчання для аналізу послідовностей. Досліджено формування кластерів даних за допомогою методу DBSCAN та чотирьох моделей вбудовувань. Проведено оцінку інформативності змінних усереднених вбудовувань, впроваджено коефіцієнт перекриття та запропоновано підхід для використання LLM безпосередньо в процесі редукції лямбда-термів. Представлено можливий варіант імплементації описаних методів для використання у компіляторах для підвищення їх продуктивності.

Висновки за результатами виконання дисертаційної роботи підкреслюють наукову новизну та практичну цінність проведених досліджень.

Список використаних джерел свідчить про те, що під час роботи було проаналізовано сучасні результати наукових досліджень.

Дисертація є завершеною науковою працею, а її оформлення відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 12.01.2022 р. № 44), та наказу Міністерства освіти і науки України від 12.01.2017 р. № 40 «Про затвердження Вимог до оформлення дисертації».

3. Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертаційна робота відповідає науковому напрямку кафедри теоретичної та прикладної інформатики факультету математики і інформатики Харківського національного університету ім. В.Н. Каразіна.

Основні результати роботи було реалізовано в Харківському національному університеті імені В. Н. Каразіна у рамках НДР «Моделювання інформаційних процесів у складних і розподілених системах» за 2021 – 2023 рр. (ДР № 0121U109183).

4. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації.

Обґрунтованість наукових положень, висновків і рекомендацій забезпечено коректним використанням широкого спектру архітектур моделей штучного інтелекту, методів імітаційного та математичного моделювання, теорії графів, та методів кластерного аналізу.

Достовірність і обґрунтованість отриманих результатів забезпечується:

- 1) застосуванням перевірених надійних моделей нейронних мереж, таких як LSTM, CNN і Transformer;
- 2) спеціально розробленим та перевіреним програмним забезпеченням для проведення експериментів;
- 3) використанням в експериментах даних, отриманих методом, що забезпечує статистичну значущість;
- 4) використанням узгоджених наборів даних у експериментах для об'єктивних порівнянь;
- 5) позитивною оцінкою результатів науковою спільнотою, що підкріплюється публікаціями та доповідями на трьох міжнародних наукових конференціях.

5. Основні наукові результати, одержані автором, та їх новизна.

Основні наукові результати та висновки дисертації пройшли апробацію під час міжнародних наукових конференцій та знайшли відображення в публікаціях у фахових та міжнародних наукових виданнях. За результатами дисертаційного дослідження опубліковано 7 наукових праць, серед яких: 2 публікації у міжнародних виданнях, які входять до наукометричних баз Scopus та Web of Science; 2 публікації у науковому виданні, включеному на дату опублікування до переліку наукових фахових видань України; 3 праці - у матеріалах і тезах конференцій.

Дисертант отримав такі **основні наукові результати**:

- 1) *вперше* розроблено параметричну випадкову однокрокову стратегію редукції, що відрізняється від існуючих можливістю налаштування ймовірності вибору редексу для скорочення. Варіація параметрів дозволяє керувати процесом редукції лямбда-термів з метою підвищення його продуктивності для конкретних підмножин термів;
- 2) *удосконалено* метод оцінки складності кроку редукції лямбда-терму, що відрізняється від існуючих використанням методів машинного

- навчання. Це дозволило підвищити точність прогнозування обчислювальної складності кроку редукції довільного лямбда-терму;
- 3) *удосконалено* експериментальне середовище емуляції лямбда-числення завдяки впровадженню процедури багатоетапної верифікації розробленого програмного забезпечення та автоматичної генерації лямбда-термів. Це дозволило забезпечити коректну реалізацію цього програмного забезпечення та створювати набори даних лямбда-термів для забезпечення теоретично необхідного покриття контрольованими конфігураціями лямбда-термів;
 - 4) *дістали подальшого розвитку* методи представлення лямбда-термів у вигляді вбудовувань (embeddings), що на відміну від існуючих представлень забезпечує використання штучних нейронних мереж архітектури типу Transformer. Це дозволило збільшити глибину аналізу коду лямбда-термів та підвищити точність екстракції характеристик, що впливають на процес редукції;
 - 5) *дістав подальшого розвитку* комплексний метод оптимізації стратегії редукції, що на відміну від існуючих поєднує моделі штучного інтелекту для оцінки часу (в кількості кроків) редукції з процедурою вибору стратегії нормалізації терму. Цей комплексний підхід є новим і має потенціал скоротити загальний час нормалізації, що суттєво сприяє оптимізації програм, реалізованих функціональними мовами програмування.

6. Практичне значення одержаних результатів.

Представлена дисертаційна робота має практичні застосування в областях функціонального програмування та оптимізації компіляторів. В ході роботи була розроблена однокрокова випадкова параметрична стратегія редукції лямбда-термів, спрямована на підвищення ефективності функціональних програм з точки зору часу виконання. У роботі пропонується метод прогнозування часу редукції конкретних редексів і загальної кількості кроків редукції з подальшою можливістю використання для автоматизації вибору стратегії редукції. Завдяки інтеграції моделей машинного навчання для прогнозування кількості кроків редукції та оцінки вимог до обчислювальних ресурсів у дослідженні пропонуються нові підходи до оптимізації всього процесу нормалізації термів у функціональних мовах програмування.

Враховуючи те, що точність моделі може мати певні обмеження, зроблені автором висновки закладають основу для майбутніх досліджень. Аналіз поведінки редукції термів і оптимізації стратегії може бути застосованим у системах, які широко використовують парадигми функціонального програмування. Здатність передбачити кількість кроків скорочення для лямбда-термів за допомогою навчених моделей може бути використана для автоматизації пошуку оптимальних стратегій редукції. Поєднуючи області теорії лямбда-числення, машинного навчання та розробки програмного забезпечення, це дослідження пов'язує функціональне програмування з більш широкими сферами науки і практики. Хоча дослідження зберігає теоретичну спрямованість, його результати мають практичні застосування, зокрема у підвищенні ефективності та результативності компіляторів функціонального програмування та оптимізації продуктивності програмного забезпечення за допомогою удосконалених стратегій редукції.

Результати дисертаційного дослідження Дейнега О. А. були впроваджені у навчальний процес Харківського національного університету імені В. Н. Каразіна, про що свідчать відповідні документи.

7. Дотримання академічної доброчесності.

За результатами аналізу дисертаційної роботи та публікацій автора порушення академічної доброчесності не виявлено. Елементи фальсифікації чи фабрикації тексту в роботі відсутні.

8. Дискусійні положення та зауваження до змісту дисертації.

Оцінюючи в цілому позитивно дисертацію Дейнега О.А., як кваліфікаційну роботу, хочу відмітити окремі недоліки та висловити певні зауваження та рекомендації. Відзначаючи новизну, важливість та повноту одержаних наукових та практичних результатів, слід відзначити й наступні недоліки:

1. В усіх розділах представлено візуалізації використаних архітектур нейронних мереж з точним описом кількості ваг. Проте в деяких випадках можна було б збільшити якість зображень архітектур використаних нейронних мереж. Дане зауваження також стосується зображень загальновідомих архітектур нейронних мереж.
2. У третьому і четвертому розділах, основними стратегіями редукції, що були використані та порівняні, є LO та RI, можливо, мало би сенс

додатково проаналізувати стратегії CallByName та CallByValue. Дані стратегії не мають значних відмінностей, проте це додало б більшої повноти та розуміння щодо проведених експериментів.

3. Не зовсім зрозумілий зв'язок між кластеризацією та класифікацією у четвертому розділі. Варто було напряду використати на даному етапі дослідження алгоритми класифікації та розглянути проблему з цього боку або навести результати таких експериментів, якщо вони були.
4. В третьому та четвертому розділах представлена робота з декількома видами вбудовувань. Беручи до уваги доволі якісний опис та наявну кількість представлених моделей, можна було б провести більш широкий огляд наявних моделей вбудовувань, та привести порівняння.
5. В третьому розділі, де представлено метод прогнозування кількості кроків редукції, наведені графіки залежності помилки від кількості кроків редукції (Рис 3.13). Треба було покращити читабельність графіків та виділити деякі з них, описати більш детально.
6. В четвертому розділі представлено таблиці інформативності змінних та таблицю коефіцієнту перекриття результатів кластерного аналізу. В даних таблицях варто було б покращити візуальну складову та додати ширший опис та пояснення знайдених залежностей та їх інтерпретації.
7. В четвертому розділі, де представлено роботу з моделлю ChatGPT, хотілося б бачити порівняльний аналіз з моделями з відкритим кодом, якщо ресурси комп'ютера дозволяють роботу з ними.
8. В четвертому розділі дослідження, де описано метод прогнозування наступного кроку редукції з використанням моделей ChatGPT, можна було б спробувати покращити використані промпти з допомогою спеціально навчених моделей LLM, що здатні генерувати якісні системні промпти.
9. Серед задач дослідження, п'ятий пункт, вказано, що оцінка кількості кроків редукції терму за певною стратегією виконується для подальшого вибору більш продуктивної стратегії для конкретного терму, варто було б розкрити такий механізм вибору детальніше.

9. Загальний висновок.

На основі детального аналізу дисертації та відповідних документів, вважаю, що дисертація «Оптимізація функціональних мов програмування на основі методів штучного інтелекту» є актуальною, завершеною науковою працею, що виконана на належному науково-теоретичному рівні з логічно та доступно викладеним матеріалом, відповідає вимогам Порядку підготовки здобувачів вищої освіти ступеня доктора філософії та доктора наук у закладах вищої освіти (наукових установах), затвердженого постановою Кабінету Міністрів України № 261 від 23 березня 2016 року (із змінами, внесеними згідно з постановами Кабінету Міністрів України № 283 від 03 квітня 2019 року, № 502 від 19 травня 2023 року та № 507 від 03 травня 2024 року), наказу Міністерства освіти і науки України № 40 від 12 січня 2017 року «Про затвердження Вимог до оформлення дисертації» (із змінами, внесеними згідно з наказом Міністерства освіти і науки України № 759 від 31 травня 2019 року) і Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України № 44 від 12 січня 2022 року (із змінами, внесеними згідно з постановами Кабінету Міністрів України № 341 від 21 березня 2022 року, № 502 від 19 травня 2023 року та № 507 від 03 травня 2024 року), а її автор, Дейнега Олександр Андрійович, заслуговує на присудження ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки» галузі знань 12 «Інформаційні технології».

Офіційний опонент,
доктор технічних наук, професор,
професор кафедри інтелектуальних комп'ютерних систем
Національного технічного університету
«Харківський політехнічний інститут»

Наталія ШАРОНОВА

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ
створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 17:22:36 12.11.2024

Назва файлу з підписом: Відгук_Шаронова_Дейнега_12_11_2024_pdf.asice
Розмір файлу з підписом: 198.1 КБ

Назва файлу без підпису: Відгук_Шаронова_Дейнега_12_11_2024_pdf.zip
Розмір файлу без підпису: 220.2 КБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: ШАРОНОВА НАТАЛІЯ ВАЛЕРІЇВНА

П.І.Б.: ШАРОНОВА НАТАЛІЯ ВАЛЕРІЇВНА

Країна: Україна

РНОКПП: 1941201842

Організація (установа): ФІЗИЧНА ОСОБА

Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 13:03:02
12.11.2024

Сертифікат виданий: КНЕДП АЦСК АТ КБ "ПРИВАТБАНК"

Серійний номер: 5E984D526F82F38F0400000062917701CF184D05

Тип носія особистого ключа: Незахищений

Алгоритм підпису: ДСТУ 4145

Тип підпису: Удосконалений

Тип контейнера: Підпис та дані в архіві (розширений) (ASiC-E)

Формат підпису: З повними даними ЦСК для перевірки (CAdES-X Long)

Сертифікат: Кваліфікований

Підписані файли: Відгук_Шаронова_Дейнега_12-11-2024.pdf

Версія від: 2024.10.11 13:00

Голові разової
спеціалізованої вченої ради
Харківського національного
університету імені В. Н. Каразіна
професору Олені ТОЛСТОЛУЗЬКІЙ
майдан Свободи 4, м. Харків, 61022

Відгук

офіційного опонента, доктора технічних наук, завідувачки кафедри систем штучного інтелекту Інституту комп'ютерних наук та інформаційних технологій Національного університету «Львівська політехніка», професора Шаховської Наталії Богданівни на дисертаційну роботу Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», подану на здобуття наукового ступеня доктора філософії з галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки.

1. Обґрунтування вибору теми дослідження.

Вибір теми дослідження обумовлений її потенціалом впливу на сферу функціонального програмування. Дослідження не тільки розглядає ключові проблеми цієї парадигми програмування, але й пропонує нові підходи до підвищення продуктивності функціональних мов програмування, їх компіляції та виконання. Функціональне програмування застосовується все частіше серед розробників для вирішення широкого спектру проблем, включно з підвищенням надійності програмного забезпечення, забезпеченням його верифікації, валідації та масштабованості. Дослідження зосереджено на методах оцінки складності редукції лямбда-термів з точки зору часу виконання кожного з кроків, що традиційно розглядаються однаковими, та загальної кількості кроків.

Представлене дослідження сприяє глибшому розумінню процесу редукції в лямбда-численні, що є основою функціонального програмування. Дослідження використовує передові методи машинного навчання для вирішення задачі прогнозування часу редукції на основі характеристик лямбда-термів. Інтеграція машинного навчання пропонує новий підхід до оптимізації компіляторів та інтерпретаторів функціональних мов програмування, що підвищує їх продуктивність та ресурсоефективність.

Підвищуючи обчислювальну продуктивність функціональних програм за допомогою удосконалених стратегій редукції, дослідження робить внесок у ширшу область розробки та оптимізації мов програмування. Дослідження характеристик, що впливають на час редукції терму, сучасними методами машинного навчання, є достатньо новим підходом в цій галузі. Продемонстровані методи відкривають можливості для автоматизованого вибору стратегій редукції, потенційно призводячи до значного

прогресу функціональних мов програмування. Зазначене свідчить про актуальність теми дисертаційного дослідження.

2. Оцінка змісту дисертації, її завершеності в цілому і оформлення.

Дисертація складається з анотацій, вступу, чотирьох розділів, висновків, списку використаних джерел та одного додатку. Загальний обсяг роботи складає 174 сторінки друкованого тексту: 143 сторінки основного тексту, 30 рисунків, 16 таблиць, списку використаних джерел із 105 найменувань.

У **вступі** обґрунтовано вибір теми дослідження, визначено мету і основні задачі, об'єкт і предмет дисертаційного дослідження, визначено наукову новизну роботи і практичне значення отриманих результатів, описано методи, які були використані у дослідженнях і особистий внесок здобувача.

У **першому розділі** представлено теоретичний аспект дослідження, з детальним описом функціональних мов програмування та лямбда-числення, а також їх взаємозв'язок. Описано переваги функціональних програм, такі як простота тестування та надійність коду, а також їх недоліки, основним з яких є низька продуктивність. Крім того, описано бібліотеку Lambda Calculus Environment, яка служить основною платформою для експериментів.

У **другому розділі** обговорюються змішані та рандомізовані стратегії редукції, а також результати експериментів та порівняння з детермінованими методами. Також пояснюється підхід для оцінки часу, витраченого на один крок редукції лямбда-терму. У розділі була розглянута концепція обчислювальної нерівнозначності редексів лямбда-термів, що є ключовою у виборі стратегії редукції.

У **третьому розділі** було проведено аналіз можливості прогнозування кількості кроків редукції лямбда-термів за заданою стратегією із застосуванням методів глибинного навчання. Також було розглянуто можливість використання вбудовувань для репрезентації різниці в редукції лямбда-термів при різних стратегіях. Для цього було протестовано чотири великих мовних моделі, для генерації вбудовувань на основі текстових представлень лямбда-термів.

У **четвертому розділі** дисертації наведено результати застосування методів неконтрольованого машинного навчання. Формування кластерів даних досліджено за допомогою методу DBSCAN і чотирьох моделей вбудовування. Після цього було оцінено інформативність усереднених змінних вбудовувань, введено коефіцієнт перекриття та запропоновано підхід для використання LLM безпосередньо в процесі редукції лямбда-терму.

Висновки за результатами виконання дисертаційної роботи підкреслюють наукову новизну та практичну цінність проведених досліджень.

Список використаних джерел свідчить про те, що під час роботи було проаналізовано сучасні результати наукових досліджень.

Дисертація є завершеною науковою працею, а її оформлення відповідає вимогам «Порядку присудження ступеня доктора філософії та скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії» (постанова Кабінету Міністрів України від 12.01.2022 р. №

44), та наказу Міністерства освіти і науки України від 12.01.2017 р. № 40 «Про затвердження Вимог до оформлення дисертації».

3. Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертаційна робота відповідає науковому напрямку кафедри теоретичної та прикладної інформатики факультету математики і інформатики Харківського національного університету ім. В.Н. Каразіна.

Основні результати роботи було реалізовано в Харківському національному університеті імені В. Н. Каразіна у рамках НДР «Моделювання інформаційних процесів у складних і розподілених системах» за 2021 – 2023 рр. (ДР № 0121U109183).

4. Ступінь обґрунтованості наукових положень, висновків і рекомендацій, сформульованих у дисертації.

Обґрунтованість наукових положень, висновків і рекомендацій забезпечено коректним використанням широкого спектру архітектур моделей штучного інтелекту, методів імітаційного та математичного моделювання, теорії графів, та методів кластерного аналізу.

Достовірність і обґрунтованість отриманих результатів забезпечується:

- 1) застосуванням апробованих та надійних архітектур штучних нейронних мереж як LSTM, CNN і Transformer;
- 2) використанням для проведення експериментів спеціально розробленого та верифікованого програмного оточення;
- 3) використанням для експериментів даних отриманих розробленим методом генерації датасетів, який забезпечує їх статистичну значимість;
- 4) використанням для порівняння результатів експериментів уніфікованих датасетів, що забезпечує об'єктивність такого порівняння;
- 5) позитивною оцінкою отриманих результатів науковою спільнотою за результатами публікацій та доповідей на трьох міжнародних наукових конференціях.

5. Основні наукові результати, одержані автором, та їх новизна.

Основні наукові результати та висновки дисертації пройшли апробацію під час міжнародних наукових конференцій та знайшли відображення в публікаціях у фахових та міжнародних наукових виданнях. За результатами дисертаційного дослідження опубліковано 7 наукових праць, серед яких: 2 публікації у міжнародних виданнях, які входять до наукометричних баз Scopus та Web of Science; 2 публікації у науковому виданні, включеному на дату опублікування до переліку наукових фахових видань України; 3 праць - у матеріалах і тезах конференцій;

Дисертант отримав наступні наукові результати:

- 1) *вперше розроблено* параметричну випадкову однокрокову стратегію редукції, що відрізняється від існуючих можливістю налаштування ймовірності вибору

редексу для скорочення. Варіація параметрів дозволяє керувати процесом редукції лямбда-термів з метою підвищення його продуктивності для конкретних підмножин термів.

- 2) *удосконалено* метод оцінки складності кроку редукції лямбда терму, що відрізняється від існуючих використанням методів машинного навчання. Це дозволило підвищити точність прогнозування обчислювальної складності кроку редукції довільного лямбда-терму.
- 3) *удосконалено* експериментальне середовище емуляції лямбда-числення завдяки впровадженню процедури багатоетапної верифікації розробленого програмного забезпечення та автоматичної генерації лямбда-термів. Це дозволило забезпечити коректну реалізацію цього програмного забезпечення та створювати набори даних лямбда-термів для забезпечення теоретично необхідного покриття контрольованими конфігураціями лямбда-термів.
- 4) *дістали подальшого розвитку* методи представлення лямбда-термів у вигляді вбудовувань (embeddings), що на відміну від існуючих представлень забезпечує використання штучних нейронних мереж архітектури типу Transformer. Це дозволило збільшити глибину аналізу коду лямбда-термів та підвищити точність екстрації характеристик, що впливають на процес редукції.
- 5) *дістав подальшого розвитку* комплексний метод оптимізації стратегії редукції, що на відміну від існуючих поєднує моделі штучного інтелекту для оцінки часу (в кількості кроків) редукції з процедурою вибору стратегії нормалізації терму. Цей комплексний підхід є новим і в має потенціал скоротити загальний час нормалізації, що суттєво сприяє оптимізації програм, реалізованих функціональними мовами програмування.

6. Практичне значення одержаних результатів.

Дослідження, проведене в представлений роботі, здебільшого теоретичне, але має практичні наслідки для функціонального програмування та оптимізації компіляторів. В ході роботи було розроблено параметричну однокрокову випадкову стратегію редукції термів лямбда-числення, яка безпосередньо впливає на підвищення ефективності функціональних програм з точки зору часу виконання. Запропоновано метод прогнозування часу редукції окремих редексів і загальної кількості кроків редукції, що допомагає автоматизувати вибір більш ефективної стратегії редукції з точки зору часу виконання. Отримані результати розширюють сферу оптимізації функціональних мов програмування за межі традиційних методів.

Попри певні обмеження точності підходу, робота створює основу для майбутніх досліджень. Розуміння поведінки процесу редукції термів і оптимізації стратегії дозволяє підвищити ефективність у системах, які широко використовують парадигми функціонального програмування.

Окрім функціонального програмування, дослідження також описує зв'язки між такими сферами, як теорія лямбда-числення, машинне навчання та розробка програмного забезпечення. Хоча дослідження зберігає теоретичну спрямованість, його

висновки мають широке практичне застосування, зокрема у покращенні ефективності компіляторів функціонального програмування та оптимізації продуктивності програмного забезпечення.

Результати дисертаційного дослідження Дейнега О. А. були впроваджені у навчальний процес Харківського національного університету імені В. Н. Каразіна.

7. Дотримання академічної доброчесності.

За результатами аналізу дисертаційної роботи та публікацій автора порушення академічної доброчесності не виявлено. Елементи фальсифікації чи фабрикації тексту в роботі відсутні.

8. Дискусійні положення та зауваження до змісту дисертації.

Відзначаючи новизну, важливість та повноту одержаних наукових та практичних результатів, слід відзначити й наступні недоліки:

1. У третьому розділі наведено діаграми класичних архітектур моделей машинного навчання для роботи з текстовим представленням даних. Оскільки дані архітектури є загальновідомими, можливо мало би сенс винести дану інформацію в додатки.
2. Сьогодні існує величезна кількість моделей машинного навчання на основі архітектури Transformer. У роботі використовуються саме модель з відомою архітектурою BERT, проте немає чіткого описання чому була обрана саме ця модель. Таке пояснення було б доречним.
3. У другому розділі представлені змішані та рандомізовані стратегії, процес підбору їх параметрів. Порівняння рандомізованих стратегій та детерміністичних можна було б провести на більшому датасеті та забезпечити більшу надійність отриманих результатів. Для змішаної стратегії, можливо, варто було б провести експерименти з використанням декількох датасетів та зробити порівняльний аналіз змішаної та детерміністичної стратегій. Це дало б більше розуміння залежності результату від датасету.
4. Опис генерації датасету в першому розділі, та подальші уточнення особливостей датасетів, використаних для експериментів варто було б зробити більш чітким. Також варто додати більш чітке описання використаних датасетів, їх розмірів та структури для кожного експерименту.
5. У роботі з рандомізованими та змішаними стратегіями не вистачає більш чітко виділеного представлення приросту продуктивності з точки зору кількості кроків порівняно з базовими стратегіями, такими як LO та RI. Можливо варто було б, додати отриману різницю у відсотках та як було сказано раніше - показати це на декількох експериментах.
6. У третьому розділі, де описано метод прогнозування кількості кроків редукції лямбда-терму в залежності від обраної стратегії, в таблиці 3.3. представлено декілька метрик, включаючи точність (ассюрасу), MAE та RMSE. Варто було б більш чітко описати роль метрики точності (ассюрасу) у задачі прогнозування

кількості кроків. Також варто було б описати зв'язок з іншими метриками та виділити основну.

7. У розділі 3 представлено гістограми розподілу довжини лямбда-термів та кількості кроків редукції (Рис. 3.7. та 3.8). На даних гістограмах варто було б збільшити якість візуалізації, надати більш широке пояснення та, можливо, візуальні позначки особливостей графіків.
8. У четвертому розділі описано результати використання підходів неінформованого навчання та представлено таблицю інформативності змінних, що отримані в результаті кластерного аналізу векторів усереднених вбудовувань (Табл. 4.2.). Можливо також можна було б провести аналіз для розуміння значення цих змінних в розрізі самого лямбда-терму, можливо деякі зміни відповідають за наявність конкретних підтермів тощо.

9. Загальні висновки щодо дисертаційної роботи.

Дисертаційна робота Дейнега Олександра Андрійовича «Оптимізація функціональних мов програмування на основі методів штучного інтелекту», представлена на здобуття наукового ступеня доктора філософії галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки є актуальною, завершеною науковою працею, що виконана на належному науково-теоретичному рівні з логічно- та доступно-викладеним матеріалом.

У роботі була розглянута важлива задача дослідження процесу редукції лямбда-термів і його залежність від структури та розміру терму за допомогою методів машинного навчання. Основним напрямком дослідження є прогнозування часу, необхідного для одного кроку редукції, і прогнозування загальної кількості кроків редукції на основі обраної стратегії. Отримані результати можуть бути застосовані для оптимізації продуктивності компіляторів та інтерпретаторів функціональних мов програмування.

Здобувач Дейнега Олександр Андрійович заслуговує на присудження йому наукового ступеня доктора філософії з галузі знань 12 – «Інформаційні технології» за спеціальністю 122 – «Комп'ютерні науки».

Офіційний опонент,
доктор технічних наук, професор,
завідувачка кафедри систем штучного інтелекту
Інституту комп'ютерних наук
та інформаційних технологій
Національного університету «Львівська політехніка».

Наталія ШАХОВСЬКА

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ
створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 15:26:15 15.11.2024

Назва файлу з підписом: Відгук_Шаховська – ЄЦП.pdf
Розмір файлу з підписом: 158.7 КБ

Перевірені файли:
Назва файлу без підпису: Відгук_Шаховська – ЄЦП.pdf
Розмір файлу без підпису: 124.8 КБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: ШАХОВСЬКА НАТАЛІЯ БОГДАНІВНА
П.І.Б.: ШАХОВСЬКА НАТАЛІЯ БОГДАНІВНА
Країна: Україна
РНОКПП: 2882502547
Організація (установа): ФІЗИЧНА ОСОБА
Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 15:26:14 15.11.2024
Сертифікат виданий: КНЕДП АЦСК АТ КБ "ПРИВАТБАНК"
Серійний номер: 5E984D526F82F38F040000002F835B019EAA0305
Алгоритм підпису: ДСТУ 4145
Тип підпису: Удосконалений
Тип контейнера: Підписаний PDF-файл (PAdES)
Формат підпису: З повними даними для перевірки (PAdES-B-LT)
Сертифікат: Кваліфікований

Версія від: 2024.10.24 15:00