Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

ФОРМАЛЬНАЯ ВЕРИФИКАЦИЯ КЛАССИФИКАЦИИ ОПРЕДЕЛИМЫХ ФУНКЦИЙ В ПРОСТО-ТИПИЗИРОВАННОМ ЛЯМБДА-ИСЧИСЛЕНИИ

Автор: Вихнин Фёдор Алексеевич _____

Направление подготовки: 01.03.02 Прикладная математика и информатика

Квалификация: Бакалавр

Руководитель ВКР: Корнеев Г.А., к.т.н. _____

Санкт-Петербург, 2023 г.

Обучающийся Вихнин Фёдор Алексеевич
Группа М34381 Факультет ИТиП

Направленность (профиль), специализация
Информатика и программирование

Консультанты:

а) Штукенберг Д. Г., магистр техничеких наук, без звания _____

ВКР принята «____» _____ 20___ г.

Оригинальность ВКР _____%

ВКР выполнена с оценкой _____

Дата защиты «15» июня 2023 г.

Секретарь ГЭК Штумпф С. А. _____

Листов хранения _____

Демонстрационных материалов/Чертежей хранения _____

Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

УТВЕРЖДАЮ

Руководитель ОП

проф., д.т.н. Парфенов В.Г. _____

«____» _____ 20___ г.

ЗАДАНИЕ
НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ

Обучающийся Вихнин Фёдор Алексеевич

Группа M34381 **Факультет** ИТиП

Квалификация: Бакалавр

Направление подготовки: 01.03.02 Прикладная математика и информатика

Направленность (профиль) образовательной программы: Информатика и программирование

Тема ВКР: Формальная верификация классификации определимых функций в просто-типизированном лямбда-исчислении

Руководитель Корнеев Г.А., к.т.н., доцент квалификационная категория ”ординарный доцент”

2 Срок сдачи студентом законченной работы до: «31» мая 2023 г.

3 Техническое задание и исходные данные к работе

Исходным моментом было рассмотрение статью 76-ого года о классе определимых функций в просто типизированном лямбда исчислении [2]. Из-за краткости и неполноты представленного доказательства в ней, связанной с отсутствие доказательств некоторых важных шагов и утверждений, заданием данной ВКР было расписать и дополнить ранее упомянутые пропуски, сформулировать чётко условия теорем и их доказательства и выполнить формальную верификацию полученного результата. Необходимость в последнем вызвано тем, что из-за человеческого фактора оба доказательства могли быть утверждены верными по ошибке.

4 Содержание выпускной квалификационной работы (перечень подлежащих разработке вопросов)

Разработанное решение на языке Arend является полной формализацией рассмотренной статьи, с введением всех необходимых для формулирования терминов, конструкций из теории типов и математического аппарата, такого как работа с множествами, и доказательством корректности всех базовых утверждений, применимых к вышеперечисленным

5 Перечень графического материала (с указанием обязательного материала)

Графические материалы и чертежи работой не предусмотрены

6 Исходные материалы и пособия

а) Оригинальная статья [2];

б) Обзорная статья с уточнением некоторых определений [3];

в) Документация по использованию языка проверки доказательств Arend [1]

7 Дата выдачи задания «01» сентября 2022 г.

Руководитель ВКР _____

Задание принял к исполнению _____ «01» сентября 2022 г.

Министерство науки и высшего образования Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

АННОТАЦИЯ
ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ

Обучающийся: Вихнин Фёдор Алексеевич

Наименование темы ВКР: Формальная верификация классификации определимых функций в просто-типизированном лямбда-исчислении

Наименование организации, в которой выполнена ВКР: Университет ИТМО

ХАРАКТЕРИСТИКА ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ

1 Цель исследования: TODO

2 Задачи, решаемые в ВКР:

a) TODO.

3 Число источников, использованных при составлении обзора: 0

4 Полное число источников, использованных в работе: 3

5 В том числе источников по годам:

Отечественных			**Иностранных**		
Последние 5 лет	От 5 до 10 лет	Более 10 лет	Последние 5 лет	От 5 до 10 лет	Более 10 лет
0	0	1	0	0	2

6 Использование информационных ресурсов Internet: да, число ресурсов: 1

7 Использование современных пакетов компьютерных программ и технологий:

Пакеты компьютерных программ и технологий	**Раздел работы**
TODOСослаться на приложение с бумажным доказательством	**??**, Приложения А, Б
TODOСослаться на репозиторий с решением	Репозиторий

8 Краткая характеристика полученных результатов

TODO

9 Гранты, полученные при выполнении работы

Никакие гранты не выделялись на разработку данного решения, все было сделано в рамках написания дипломной работы

10 Наличие публикаций и выступлений на конференциях по теме выпускной работы

Никакие публикации и выступления на конференциях не проводились

Обучающийся Вихнин Ф.А. _____

Руководитель ВКР Корнеев Г.А. _____

«____» _____ 20___ г.

СОДЕРЖАНИЕ

СПИСОК ТЕРМИНОВ

а) Выражение
б) Тип
в) Типизация выражения
г) Альфа-эквивалетность
д) Бета-редукция
е) Бета-экивалетность
ж) Свободная переменная
и) Нормальная форма
к) Нумерал

ВВЕДЕНИЕ

Теория типов, как раздел математики, является важной частью не только интереса учёных в области компетенции но и также важный составляющий мира практического программирования, так как модели, используемые и разрабатываем в рамках данной в рамках данной дисциплины, используются в качестве архитектур ныне существующих языков программирования, технических решений либо же объясняют свойства и как-то формализуют системы, которые используются в большинстве языков программирования. В свою очередь просто типизированное лямбда исчисление является одной из самых простых, но не менее важных, моделей выражений в языках программирования. Утверждения, которые доказываются по отношению к ней, имеет большое как теоретическое, так и практическое значение. Потому есть потребность, также как любой науке, чтобы эти утверждения были заведомо корректные, особенно в случае, если выводы, которые делаются, дают какое-то большое представление об области. В свою очередь, так как все доказательства делают людьми нельзя исключать человеческий фактор. Существует множество примеров, в которых доказательства многие которые многие годы поддерживались и считались по итогу оказывались неверными и по ошибке предполагались таковыми, потому что не были рассмотрены какие-то случаи или проверяющие специалисты не обнаруживали каких-то неувязок и неточностей в оригинальном доказательстве, что могло приводить либо к появлению новых теорий и новых решений, либо же ставило под вопрос выводы всех последующих в данной области работ, так как стартовая точка оказывалась неверной. В рамках данной ВКР мы рассматриваем как раз одно из таких утверждений, заключающаяся в том, что все функции результирующие из полиномов являются полиномами. Так утверждение является достаточно важным и большим результатом в данной области очень не хотелось бы столкнуться с его некорректностью. Доказательство оригинальной статьи нельзя считать полным, так как оно обладает некоторыми допущениями и не представляет доказательства каждого шага, а также очень коротко поясняет все свои выкладки из чего следует не точность как понимание заложенного автором, так и подозрение вне точности представленных им рассуждений.

0.1. Актуальность работы

TODO

0.2. Цели и задачи

TODO

В рамках данной ВКР ставились три цели

Первая цель: изучение и проверка оригинальной статьи на корректность, её уточнение в случае необходимости и расширение новыми доказательствами. По итогу работу пришлось дополнить четырьмя новыми утверждениями доказывающими некоторые шаги в оригинальной статье или же уточняющая существующее решение.

Вторая цель: перенос всех полученных утверждений в язык проверки доказательств. То есть получение сертификации корректности всех этапов решения.

Третья цель: расширение кодовой базы в области теории типов для дальнейшей для возможности изучения, проверки доказательств и формулирования новых утверждений сразу же пользуясь данным инструментом.

0.3. Новизна работы

TODO

0.4. Практическое значение работы

TODO

Практическим значением работы считается представление большой кодовой базы в области теории типов именно на языке аренд, как новый языка для доказательств, использующего более современные теории, по сравнению со своими предшественниками. Что влечёт за собой не только развитие самого языка, но и самой области формальных верификаций утверждений про языки и математические теории. В дальнейшем такое же бурное развитии данной области к развитию компиляторов, делая их более умными и способными на большие возможности в проверке полученного кода, также развитие автоматической проверки кодовых решений на баги, уязвимости и другие трудности, также вытекающие из природы написание кода людьми.

Также считаю важным отметить, что результатом данной работы является улучшение представления о возможности формализации и представления нынешней теории в области просто-типизированного лямбда исчисления и математической логики. Например, в рамках работы были рассмотрены различные способы представления выражение как таковых и их

альфа-эквивалентности (содержательной одинаковости), благодаря полученному опыту были сделаны выводы об эффективности и простоте использования одних, таких как аннотация Де Брауна, и неудобности более классических, как например выражение с именными переменными.

0.5. Краткое описание

TODO

ГЛАВА 1. ИЗУЧЕНИЕ ОРИГИНАЛЬНОЙ СТАТЬИ И ЕЁ ДОПОЛНЕНИЕ

TODO

1.1. Предметная область

TODO

1.2. Описание статьи

Оригинальная статья посвящена классификации определимых функций в просто-типизированном лямбда исчислении. Первое, что делает автор, так это упоминает о не типизированном лямбда исчислении, и об определимости функций в нем, и задается вопросом о таковых в типизированном случае. Стоит отметить, что статья носит очень неформальный характер, потому о каком точно определении речь не сразу ясно, как будет показать в статье 2, детали описания имеют непосредственное влияние на возможные рассуждения.

В краткой форме упоминается классической определение строго определимых функций с уточнением типа равенства, который использовался. Это равенство по отношению beta редукции, то есть ”структурное”равенство, когда за конечное количество шагов редукции можно прийти к одной и той же форме выражения, с точностью до имен переменых (отношения альфа-эквивалентности). Важно отметить, что именно о структурном равенстве идет речь, а не о ”вычислительном когда гарантируется лишь одинаковое поведение функций на одинаковых данных (ню-экивалетность, ссылка на статью 3).

И далее автор дает определение ”расширенных полиномов класса функций, замкнутого по композиции:

а) сложения
б) умножения
в) числовых констант
г) функции проверки на ноль (”иф-зиро”), которая принимает три аргумента и если первый аргумент ноль, то возвращает второй, иначе третий

Далее следует доказательство того, что только функция из такого класса, может быть строго определима. И теперь перейдем к рассмотрению приведенного доказательства и какие недостатки имел оригинал.

TODO

1.3. Анализ оригинального доказательства

Автор начинает рассуждает с того, что класс чистых функций (то есть функций, результат которых зависит только от переданных значений) или же замкнутых форм (то есть выражений, которые не имею свободных переменных (ссылка на определение)) замкнут по композиции. Данное утверждение не доказывается, но его можно считать очевидным.

Далее автор упоминает, что все расширенные полиномы являются определимыми функциями, хоть этот факт тоже им остается не доказанным. Данное утверждение является крайне важным, так как таки образом автор подчеркивает, что множество таких функций является подмножеством определимых. И далее как он покажется зеркальное утверждение (что определимые функции являются подмножеством расширенных полиномов), то он сможет утверждать равенство данных классов.

Пользуясь тем, что в расширенных полиномах, добавлена функция иф-зиро, он говорит, что также все функции результирующие из полиномов путем добавления или удаления аргумента - тоже представимы, так как из-за конечности числа аргументов, можно рассмотреть конечное количество случаев с рассмотрением значений переменных на равенство нулю, что выливается в построение дерева разбора случаев из иф-зиро по каждой переменной.

В таком случае остается убедиться лишь в том, что иных функций нет.

Для этого в качестве примера, он рассматривает двуместную функцию t (для любой другой мерности, доказательство аналогичное), такую, что она определима, а значит у нее есть выражение замкнутой формы, типизируемое, как $\nu \to \nu \to \nu$, то есть как функцию от двух переменных типа ν - типа ассоциированного с нумералами - местным представлением чисел.

Взяв в качестве аргументов для данной функции числа n, m и им соответствующие выражения F, G, а также рассмотрев функцию $\alpha : \tau \to \tau$ как третий аргумент для выражения t, он применяет их к выражению, берет от полученного нормальную форму и смотрит на полученных результат. Так как равенство здесь рассматривается как бета-экивалетность, то приведение к нормальной форме выражения является корректным шагом, сохраняющего справедливость утверждений об оригинальном выражении.

Так как изначально бралось выражение, которое типизируется, то нормальная форма существует и она единственна, по теореме Черча-Россера (ссылка на источник).

Рассмотрение именно нормальной формы полученного выражения, позволяет избавиться от некоторых затрудняющих случаев при рассмотрении под-термов выражения. Если быть точным: случай рассмотрения аппликации, так как в случае нормальной формы, левая часть не может быть лямбда-абстракцией, иначе бы можно было применить бета-редукцию.

Далее идет разбор возможных под-термов, полученного выражения, с учетом их типизации.

Первое, что утверждается, это то, что любой подтерм имеет один из трех типов:

а) τ

б) $\tau \to \tau$

в) $\nu = (\tau \to \tau) \to \tau \to \tau$

Данное заявление никак не доказывается, и является первым таким утверждением, крайне мала очевидность правдивости которых. В дальнейшем будет рассмотрена данная проблема.

Далее утверждается, тоже бездоказательно, что единственным подтермами типа ν являются константы F, G. И разбираются возможные случаи подтермов, которые могут иметь тип $\tau \to \tau$.

Опуская неочевидность существования только таких случаев, автор приближает доказательство к концу, воспользовавшись математической индукцией по структуре выражения. А точнее, он хочет показать, что любой подтерм данного выражения с типом $\tau \to \tau$ может быть средуцирован до выражения вида $\lambda y.\alpha^{P(n,\,m)}z$, где P - функция из класса расширенных полиномов, а z - любая переменная. В дальнейшем такое выражение будет называться ”константным полиномом”. Показывает он это тем, что рассматривает каждый ранее указанный разбор такого выражения как шаг индукции, за базу взяв случай равенства выражения с ранее упомянутой α.

По итогу доказав то, что любой такой подтерм является константным полиномом, он, ссылаясь на то, что искомое выражение $tFG\alpha$ тоже имеет такой тип, получает, что функция редуцируется в $\lambda y.\alpha^{P(n,m)}y$ - что является нумералом от $P(n, m)$, а значит полученная функция - расширенный полином.

TODO

1.3.1. Рассмотрение пропущенных шагов доказательства

Как было указано в предыдущей главе, автор пользуется перечнем недоказанных им и неочевидных также утверждений, на основе которых и строит свои рассуждения, и без которых доказательство не представляется возможным, а именно:

а) Любой подтерм обладает одним из перечисленных типов: $\tau, \tau \to \tau, \nu$

б) Любой подтерм типа ν - это F или G.

в) Любой подтерм типа $\tau \to \tau$ - это либо

1) α
2) Fs или Gs, где s - также типизируется как $\tau \to \tau$
3) $\lambda y.s_1(s_2(\cdots s_k(z)\cdots))$, аналогично $\forall i.\ s_i$ имеет тип $\tau \to \tau$

Перечисленные шаги не являются выводами из каких-то теорем или не являются результатами каких-то работ, по-крайней мере нет ни единого упоминания о таковых, а с учетом специфики области и новизны такого класса функций, на момент написания работы, и, следовательно, рассуждений о структуре выражений в таком ключе и с такими типами, то закономерно заключить, что все перечисленные утверждения, особенно с учетом их содержательности, были оставлены автором непроверенными и требуют быть доказанными, дабы можно было считать искомое доказательство верным.

тут странно. TODO

1.3.2. Необходимость в ведении новых утверждений

TODO

Выводы по главе 1

В данной главе была рассмотрена оригинальная статья, значимость ее результата для области и досконально разобрано предлагаемое доказательство.

В ходе рассмотрения доказательства было выявлено несколько мест, где автор формулировал для дальнейшего использования утверждения, доказательство которых не предоставлялось ни самим автором, ни в качестве ссылки на другие работы в данной области. Именно данный нюанс и побудил рассмотрение доказательство подробнее, дополнить его недостающим доказательствами и утверждениями и произвести процесс полной верификации полученного результата.

ГЛАВА 2. МЕТОДОЛОГИЯ РЕШЕНИЯ

TODO

2.1. Анализ необходимой теоретической базы

TODO

2.2. Этапы решения

TODO

Выводы по главе 2

ГЛАВА 3. ФОРМУЛИРОВАНИЯ ДОКАЗАТЕЛЬСТВА НА ЯЗЫКЕ AREND

TODO

3.1. Мотивация выбора данного языка

TODO

3.2. Сложности во время реализации

TODO

3.2.1. Альфа эквивалентность

TODO

3.2.2. Операции над множествами

TODO

3.3. Практическое применение

TODO

Выводы по главе 3

В конце каждой главы желательно делать выводы. Вывод по данной главе — нумерация работает корректно, ура!

ЗАКЛЮЧЕНИЕ

В данном разделе размещается заключение.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1 Arend Theorem Prover [Электронный ресурс]. — URL: https://arend-lang.github.io/documentation/.

2 *Schwichtenberg H.* Definierbare Funktionen im λ-Kalkul mit Typen. // Archiv für mathematische Logik und Grundlagenforschung, No. 17. — 1976. — S. 113–114. — URL: https://epub.ub.uni-muenchen.de/4273/1/10.pdf.

3 *Zakrzewski M.* Definable functions in the simply typed lambda-calculus. — 2007. — DOI: 10.48550/arXiv.cs/0701022.

ПРИЛОЖЕНИЕ А. МАТЕМАТИЧЕСКИЕ ВЫКЛАДКИ

ПРИЛОЖЕНИЕ Б. РЕПОЗИТОРИЙ С КОДОВЫМ РЕШЕНИЕМ