

Санкт-Петербургский государственный политехнический университет
Институт информационных технологий и управления
Кафедра распределенных вычислений и компьютерных сетей



Работа допущена к защите

Зав. кафедрой

_____ Ю. Г. Карпов

" ____ " _____ 2014г.

ВЫПУСКНАЯ РАБОТА БАКАЛАВРА

Тема: *Разработка и реализация модульной системы проверки и
вычисления типов*

Направление: *123456 — номер и название направления*

Выполнил студент гр. 43507

Руководитель, ст. преп.

М. А. Буряков

Д. А. Тимофеев

Санкт-Петербург
2014

Утверждаю
Зав. кафедрой
_____ Ю. Г. Карпов
" ____ " _____ 2014г.

ЗАДАНИЕ
на дипломную работу
студенту М. А. Бурякову

1. Тема: *Разработка и реализация модульной системы проверки и вычисления типов*
(утверждена распоряжением по институту от _____ № _____)
2. Срок сдачи работы.
3. Исходные данные к проекту (работе).
4. Содержание расчетно-пояснительной записки (перечень подлежащих разработке вопросов).
5. Перечень графического материала с точным указанием обязательных чертежей.
6. Консультанты по проекту (с указанием относящегося к ним разделов проекта, работы) .

Дата выдачи задания: _____ г.

Руководитель: _____ ст. преп. Д. А. Тимофеев

Задание принял к исполнению: _____ М. А. Буряков

Реферат

С. 37

Краткая характеристика всего документа, основной результат работы,
список ключевых слов

Abstract

37 pages

Brief description of the work, main result, keywords

Оглавление

Список обозначений	7
Список терминов	8
Введение	10
1 Системы типов	15
2 Используемые технологии	18
2.1 Среда MPS	18
3 Язык описания систем типов	20
3.1 Парадигма языка	20
3.2 Репозиторий типов	21
3.2.1 Аспекты репозитория	21
3.2.2 Многократная запись в репозиторий	23
3.3 Правила типизации	23
3.3.1 Аннотации правил	24
3.3.2 Подбор правил	25
3.3.3 Автоматическая генерация аннотаций	26
3.4 Операции над типами	26
3.5 Типовые переменные	27
4 Правила вывода типов для языка Haskell	29
4.1 Объявление алгебраических типов	29
4.2 Лямбда-выражение	30
4.3 Правило для всех выражений	31
4.4 Применение функции к аргументу	31

4.5	Ссылка на переменную лямбда-выражения	31
4.6	Ссылка на объявление выражения или конструктора	32
4.7	Объявление выражения	32
4.8	Выражение let	33
4.9	Выражение case и сопоставление с образцом	33
5	Описание системы типов языка Java	35
	Заключение	36

Список обозначений

АБГШ	аддитивный белый гауссовский шум
AST	абстрактное синтаксическое дерево
MPS	Metaprogramming system
GADT	Generalized algebraic datatypes

Список терминов

Абстрактное синтаксическое дерево (abstract syntax tree, AST) -

Приписывание типов - сопоставление некоторым вершинам AST корректно составленных типов.

Проверка типов - алгоритмическая процедура, анализирующая синтаксически корректное AST программы с приписанными типами и определяющая, является ли AST с приписанными типами корректным по типам.

Корректность по типам AST с приписанными типами - результат работы проверки типов.

Типизируемость AST (типизируемость программы) - существование такого приписывания типов, что AST с данным приписыванием типов корректно по типам.

Единственность типизации AST - единственность такого приписывания типов, что AST с данным приписыванием типов корректно по типам.

Единственность типов - свойство системы типов, заключающееся в том, что любое синтаксически корректное AST единственно типизируемо.

Частичная проверка типов - проверка типов для части AST с приписанными типами. Так как часть AST не является синтаксически корректным AST, требование синтаксической корректности снимается. Если часть AST с приписанными типами не прошла частичную проверку типов, то целое AST с такими же приписанными типами не должно пройти полную проверку типов.

Вычисление типов - алгоритмическая процедура, производящая приписывание типов. Может завершаться неудачей.

Система типов - множество корректно составленных типов, процедура вычисления типов, процедура проверки типов.

Язык программирования - определяет множество синтаксически кор-

ректных AST и систему типов.

Гарантированный вывод типов – алгоритм вычисления типов, обладающий следующим свойством. Если он завершает работу, то AST с написанными типами корректно по типам, а если не завершает, то AST не типизируемо.

Восстановление типов – то же, что и гарантированный вывод типов, в случае единственности типов.

Аннотация типа – упоминание типа в AST программы. Если типы записываются в виде деревьев, то аннотацией типа можно считать включение типа в AST программы в качестве поддеревя. Аннотации типов, как правило, непосредственно указывают типы вершин AST.

Переменная типа (типовая переменная) –

Введение

Большая часть современных языков программирования имеют статическую систему типов для классификации конструкций, использующихся при описании программы (переменные, выражения, функции), по типам данных, которые будут соответствовать этим конструкциям в момент выполнения программы. Статическая система типов может рассматриваться как ограничение на абстрактное синтаксическое дерево программы [1]. Проверка и вычисление типов осуществляется путем обхода этого дерева, в ходе которого каждому узлу дерева приписывается некоторый тип. При этом типы некоторых конструкций задаются пользователем явно, а другие должны быть вычислены автоматически. Алгоритмы, по которым производится вычисление и проверка типов, различны для разных языков и задаются, как правило, в описании языка. Но среди различных алгоритмов вычисления и проверки типов многие алгоритмы отличаются лишь в деталях или имеют похожие элементы. Поэтому при создании инструментов, предназначенных для работы с различными языками, например, интегрированных среды разработки, возникает потребность в модульной системе, позволяющей описывать различные алгоритмы вывода типов, выделяя общие элементы для использования в различных алгоритмах.

Традиционно алгоритмы вычисления и проверки типов реализуются отдельно в каждом компиляторе или среде разработки для того или иного языка. Потребность в объединении различных алгоритмов появилась с появлением платформ для разработки языков (Jetbrains MPS, Xtext, Spoofax и др.). Некоторые платформы для разработки языков включают инструменты для описания систем типов. Подход, предлагаемый Jetbrains MPS для вычисления типов, основан на создании переменной типа для каждой типизируемой конструкции программы и построении системы уравнений и неравенств относительно этих переменных. После то-

го, как система уравнений и неравенств будет построена, встроенный решатель пытается найти решение этой системы, которое и будет представлять вычисленные типы для соответствующих конструкций программы. Недостатком данного подхода является неопределённость выбора решения, удовлетворяющего требуемым уравнениям и неравенствам. Если уравнения и неравенства составлены так, что их решение не единственно, то в результате работы алгоритма будет выбрано одно произвольное решение, причём при выборе используются некоторые недетерминированные эвристические соображения, в результате чего ответ может изменяться от запуска к запуску. Для получения корректного результата разработчику языка требуется дополнить уравнения и неравенства дополнительными условиями, чтобы устранить неопределённости. Однако практика показывает, что для языков, система типов которых не была изначально приспособлена для метода решения уравнений и неравенств (например, Java), записать уравнения и неравенства так, чтобы решение было однозначным, является достаточно сложной задачей и удаётся далеко не всегда.

Требуется разработать систему библиотечных или языковых конструкций (при необходимости, язык или набор языков), позволяющий описывать различные алгоритмы вычисления и проверки типов в виде, подобном их неформальному описанию в виде правил. Правила, описываемые в этой системе, должны быть по возможности локальными, то есть каждое правило должно применяться к небольшому фрагменту AST, содержащему определённую конструкцию языка. Правила должны быть комбинируемыми, то есть при добавлении к языку новой конструкции изменение системы типов должно сводиться к добавлению одного или нескольких правил. Правила должны быть понятными по отдельности, то есть записанными на наглядном языке, и каждое правило должно представлять собой некоторую смысловую единицу, описанную в спецификации языка.

Для иллюстрации разрабатываемой системы планируется описать в ней алгоритмы вывода типов языка Haskell 98 (алгоритм Хиндли-Милнера с *let*-полиморфизмом) и языка Java 6. Данный выбор языков продиктован следующими причинами: алгоритм Хиндли-Милнера является наиболее известным алгоритмом вывода типов с помощью уравнений, а язык Java, в частности его статическая система типов, снабжён качественной спецификацией, структурное соответствие с которой может служить критерием наглядности записи правил. Также важным является тот

факт, что конструкции языка Java уже реализованы в среде MPS, и также в среде MPS разработано несколько расширений языка Java.

Правила типизации обычно описываются на некотором в той или иной степени формальном языке. Семантика такого языка достаточно декларативна, что не позволяет её использовать непосредственно для построения алгоритма, находящего корректную типизацию. Заметим, что, согласно определению [tapl, 8.2.1], типизация является наименьшим отношением, удовлетворяющим правилам. Такое определение неудобно для построения алгоритма, поэтому алгоритм вывода типов простого типизированного лямбда-исчисления основывается на лемме инверсии [tapl, 8.2.2; 9.3.1], которая позволяет для конкретного терма вычислить его тип. Поэтому правила, записанные декларативно, будем называть условиями или ограничениями типизации, а правила, позволяющие построить алгоритм приписывания типов (в данном случае основанные на лемме инверсии) — правилами вывода типов.

Замечание 1.1. Под правилами вывода типов будем понимать небольшие, локально действующие правила, которые определены для языка заранее, и их конечное число (так же как конечно число аксиом – аксиомопорождающая схема является одной аксиомой), приблизительно по числу значимых конструкций языка. Целый алгоритм вывода типов правилом не является, алгоритм реализуется в последовательном (лучше даже параллельном) применении определённым образом трактуемых правил.

Получение правил вывода типов из условий типизации является весьма сложным процессом, и в общем случае, выполняется вручную. Поэтому не будем ставить задачу автоматизировать данный процесс, хотя его автоматизация для конкретных классов систем типов вполне возможна и представляет практический интерес. Предположим, что разработчик системы типов будет её записывать уже в виде правил вывода типов. Тогда встанёт вопрос о том, как будут выглядеть эти правила вывода типов. В [tapl] приводятся примеры правил вывода для систем типов, не содержащих *type inference*. Однако *type inference* привносит некоторые сложности в язык описания правил вывода типов, особенно в сочетании с другими возможностями системы типов.

В частности, для систем типов, осуществляющих *type inference*, не подходят правила, берущие тип одних термов и вычисляющие на его основе тип других термов. Вместо этого используются правила, создающие уравнения (в общем случае это могут быть произвольные ограничения, например, для *bounded quantification* появятся неравенства). Уравнения на-

капливаются в специальном пуле, а затем решаются с помощью определённого алгоритма.

Таким образом, можно выделить два подхода к описанию правил вывода типов: через вычисление, когда типы одних термов объявляются функционально зависящими от типов других термов, и через ограничения, когда правила вывода лишь создают уравнения, которые впоследствии уравнения решаются специальным алгоритмом.

При реализации и описании системы типов конкретного языка выбирается тот или иной подход в зависимости от особенностей языка. Но, когда разрабатывается система типов для встраивания в *language framework*, заранее неизвестно, какого вида язык будет описан, поэтому требуется остановиться на каком-то одном подходе, уточнить его и все возможности системы типов реализовать через него.

Замечание 1.2. Можно, конечно, предложить две системы типов, каждая из которых будет использовать свой подход (язык правил вывода), но в таком случае возникнет вопрос их взаимодействия. Более того, даже в этом случае появится желание на каждом языке правил вывода выразить как можно больше возможностей системы типов, что и предлагается сделать в данной работе на примере одного языка правил вывода.

Подход через вычисление весьма прост в реализации, но есть мнение, что он не подходит для систем, осуществляющих *type inference*. У подхода к выводу типов через решение уравнений имеется ряд сложностей. Во-первых, правила, удобно записываемые через вычисления, весьма неудобно выглядят, будучи переписанными в виде уравнений, а зачастую они и вовсе не могут быть переписаны. При записи уравнениями нельзя взять тип той или иной вершины и произвести с ним нетривиальное вычисление, потому что в момент создания уравнений тип неизвестен, а в какой этап алгоритма решения уравнений встроить это вычисление, не очень понятно. Был предложен приём, позволяющий откладывать это вычисление до окончательного вычисления типа некоторой вершины, но использование этого приёма откладывает типизацию некоторых других вершин, что приводит к необходимости создавать уравнения в процессе решения уравнений, а это чревато непредсказуемыми ошибками. Во-вторых, есть проблемы с самим алгоритмом решения уравнений. Когда ограничения представляют собой только уравнения, успешно используется алгоритм унификации, но когда возникают неравенства, предложенный на данный момент алгоритм становится недетерминированным. Добавление произвольных отношений между типами только усугубит недетерминированность,

так как система отношений может иметь много решений, и нет способа однозначно выбрать то или иное решение, удовлетворяющее условию.

Задачей данной работы является уточнение языка описания правил вывода типов, использующего в качестве основы парадигму вычислений, и адаптация языка для того, чтобы на нём можно было выразить системы типов, выполняющие *type inference*. Одновременно с уточнением языка правил также необходимо определить способ применения правил (некоторую виртуальную машину), иначе семантика языка описания правил будет не вполне ясна.

В качестве примера систем типов, осуществляющих *type inference*, планируется использовать упрощённые или модифицированные варианты языков Haskell и Java. В случае успешного применения предложенных алгоритмов к данным языкам возможно рассмотрение систем высших порядков или сложных типсистемных расширений объектных языков.

Глава 1

Системы типов

2. Системы типов в языках программирования.

Понятие типа и системы типов можно определить по-разному. Мы будем рассматривать систему типов как синтаксический метод нахождения в исходной программе некоторых видов ошибок. Так как метод синтаксический, то оперирует он представлением программы в виде абстрактного синтаксического дерева (AST). Метод этот состоит из двух этапов: сначала каждой вершине AST программы приписывается некоторый тип, затем производится проверка корректности программы вместе с приписанными типами.

Замечание. Как правило, тип требуется не каждой вершине, а только выражениям языка. В дальнейшем будем это опускать, и, говоря, что вершине приписывается тип, будем подразумевать вершины, соответствующие выражениям.

Замечание. Этап вычисления типов теоретически является необязательным – программиста может все типы аннотировать явно. Но это оказывается крайне неудобно, поэтому желательно свести к минимуму необходимое количество аннотаций. В идеальном случае программист может совсем не указывать аннотаций типов, а некоторый алгоритм либо подберёт такие типы, чтобы типизированная программа прошла проверку типов, либо докажет, что такие типы подобрать невозможно. На практике такой идеальной ситуации гарантированного вывода типов без аннотаций удаётся достичь лишь для отдельных классов систем типов, а в остальных случаях либо от программиста требуются некоторые обязательные аннотации, позволяющие гарантировать алгоритмическую разрешимость

дальнейшего гарантированного вывода типов, либо вывод типов проводится без гарантии того, что если он не удался, то невозможно подобрать корректные типы.

Сложность и эффективность алгоритмов, работающих как на первом этапе вычисления типов, так и на втором этапе проверки типов, может варьироваться в зависимости от возможностей, которые предоставляет система типов.

2.1. Простой вариант системы типов содержит конечное число типов, которые встроены в систему типов. Листья AST являются либо константами, либо ссылками на переменные – и те, и другие имеют заранее известный тип, так как тип констант виден из вида литерала, а тип переменных указывается при объявлении переменной. В такой системе типов вычисление типов производится снизу вверх по дереву, то есть тип любой вершины может быть вычислен, если известны типы нижележащих вершин. Проверка типов также происходит просто, так как все требования локальны и могут быть описаны таблично.

2.2. Если язык программирования разрешает пользователю вводить новые типы, то вместе с новыми типами должны быть объявлены конструкторы этого типа и операции с выражениями этого типа. Возможны ситуации, когда конструкторы и операции задаются неявно при каждом объявлении типа. Может быть и так, что конструктор не требуется – достаточно лишь объявиться переменную этого типа.

Как правило, добавление в язык возможности объявления пользовательских типов не сильно влияет на алгоритмы вычисления и проверки типов, так как можно перед началом вычисления типов обойти все объявления новых типов, совершить некоторые предварительные вычисления, а затем рассматривать эти типы как встроенные в язык.

2.3. Подтипы в их наиболее простом виде добавляют одно отношение частичного порядка над типами – отношение подтипа. При этом правила проверки типов модифицируются таким образом, что в любом месте программы, где корректно выражение некоторого типа, корректным становится и выражения всех его подтипов. Такой подход к подтипам весьма интересен в теоретическом плане, однако никак не способствует созданию алгоритма вычисления типов. Поэтому в большинстве языков программирования, содержащих подтипы, алгоритм вычисления типов задаётся отдельно.

#Неразрешимость системы типов Java. Пример некорректной работы `javac`. `F-bounded quantification` и кому это нужно

(github.com/ohnosequences/ohnosequences.github.com/blob/blog/post/typed/graphs,
posts/2014-05-18-typed-property-graphs-in-java.markdown) Enum #

Глава 2

Используемые технологии

2.1 Среда MPS

Для практического построения моделей систем типов планируется использовать такой language framework как JetBrains MPS. Удобство его использования для этой цели заключается в том, что он позволяет легко описывать языки, а также предоставляет некоторый высокоуровневый API для работы с кодом, написанном на этих языках. MPS уже содержит язык описания правил вывода типов, однако язык этот основан на создании уравнений, а следовательно, достаточно ограничен.

Использование MPS подсказывает некоторые особенности машины системы типов. Во-первых, типы присваиваются не терму, оторванному от контекста, а конкретной вершине AST. Это вполне отвечает практической необходимости, потому как возможны системы типов, где одно и то же поддерево AST, повторенное в разных местах дерева, будет иметь в них различный тип. Во-вторых, сами типы также записываются в виде AST. В случае простых типов это не имеет особого значения, но уже появление параметризованных типов (которые ещё и могут встречаться в тексте программы) выявляет общую природу кода на языке и типов этого языка.

Среда MPS позволяет использовать в одной программе разные языки, который, в общем случае, могут быть разработаны независимо. Для совместимости языков, в числе прочего, требуется совместимость систем типов. Для этого требуется, как минимум, существование движка системы

типов, позволяющего разработчикам языка описывать правила вычисления типов. Таким образом, нужно прояснить набор базовых понятий, которые нужны для описания правил системы типов, причём этот набор должен быть достаточно богат, чтобы описать как Haskell-like, так и ML-like и Java-like системы типов. При этом следует оставить как можно меньше универсальных, основополагающих понятий, а остальные выразить через них. При этом для описания систем типов разных классов можно создать несколько удобных языков, генерируемых в минималистичный язык/API движка типсистемы.

Глава 3

Язык описания систем типов

3.1 Парадигма языка

Языки, допускающие лишь некоторый ограниченный набор операций, не позволяющий выразить достаточно сложные вычисления, подходят лишь для заранее очерченного круга задач. Поэтому логично составлять язык для виртуальной машины, выводящей типы, так, чтобы позволить пользователю производить произвольные вычисления с типами и другими объектами. Для этих целей язык описания систем типов допускает включение в определённых местах произвольных фрагментов кода на языке общего назначения (Java с расширениями).

Альтернативным вариантом мог бы являться полностью декларативный подход к описанию языков, при котором не допускается использование тьюринг-полного императивного языка для описания вычислений. Но такой полностью декларативный подход к описанию систем типов имеет недостатки.

Во-первых, системы типов большинства разрабатываемых предметно-ориентированных языков описываются в терминах последовательных вычислений. В результате такие, как правило, простые системы типов будут требовать значительных усилий, чтобы их описать на декларативном языке. Перевод в обратную сторону — из декларативного описания в спецификации языка в описание с использованием

императивного языка — представляется менее затратным относительно сложности самой системы типов, так как та или иная декларативная семантика спецификации системы типов используется для языков общего назначения с достаточно сложной системой типов, при этом разработчик, взявшийся за непростую задачу описания языка общего назначения в универсальной среде описания языков, в любом случае будет тратить достаточно много времени на осмысление семантики описания системы типов этого языка.

Во-вторых, если язык описаний систем типов не будет тьюринг-полным, то в какой-то момент найдётся система типов, требующая таких вычислений, которые невозможно выразить на этом языке. Если же этот язык сможет описать сколь угодно сложные вычисления, то этот язык окажется достаточно сложным по структуре и семантике, и будет весьма непростым для изучения. Поэтому в данной работе принято решение использовать язык общего назначения (Java) для описания сколь угодно сложных вычислений, и специально разработанный предметно-ориентированный язык для оперирования понятиями, специфичными для задачи вывода и проверки типов.

3.2 Репозиторий типов

Язык для виртуальной машины, выводящей типы, позволяет производить произвольные вычисления с типами и другими объектами. Логично потребовать, чтобы все вычисления, производимые в процессе вывода и проверки типов, не имели побочных эффектов. Это означает, что все обращения к типизируемому коду, обращения к состоянию машины, выводящей типы, и изменения состояния машины производятся только через специальные команды. Состояние машины содержит типы, присвоенные вершинам. Они хранятся в виде репозитория типов. Репозиторий типов представляет собой хранилище, сопоставляющее каждой вершине абстрактного синтаксического дерева (AST) вычисленный для неё тип. Взаимодействие с репозиторием типов осуществляется через две операции: *чтение типа* и *запись типа*.

3.2.1 Аспекты репозитория

Кроме непосредственно типов, может понадобиться хранить результаты промежуточных вычислений. Их тоже удобно хранить привязанными к

определённой вершине абстрактного синтаксического дерева, так же как и типы, в репозитории. Для этой цели введена возможность присваивать одной и той же вершине абстрактного синтаксического дерева одновременно несколько «типов», различая их по специальным меткам, *аспектам репозитория*.

Аспекты репозитория объявляются отдельно. Чтение из репозитория и запись из репозитория производится с указанием аспекта репозитория, по которому должно быть произведено чтение или запись, и вершины AST, к которой привязана читаемая или привязывается записанная информация. В большинстве практических случаев по одному из аспектов репозитория хранятся собственно типы выражений языка, а другие аспекты объявляются для вспомогательных целей. Примером осмысленного дополнительного аспекта может быть аспект, по которому для некоторой вершины записывается максимальный допустимый тип выражения, которое могло бы стоять на месте этой вершины. Такой аспект чрезвычайно полезен для автоматического дополнения кода в интегрированной среде разработки.

Также множественность аспектов полезна и с точки зрения совместимости. Если различные разработчики независимо друг от друга создали системы типов для одного языка, и есть необходимость использовать обе эти системы, то они могут работать, не мешая друг другу, если будут использовать различные аспекты репозитория. При этом каждая система получит как бы свой репозиторий, и эти «репозитории» можно считать независимыми.

Особенно полезно одновременное использование различных систем типов может быть если в рамках модели системы типов осуществляются различные проверки, не относящиеся непосредственно к типам так, как они описаны в спецификации языка, но, подобно системам типов, производящие более сложные статические проверки, в том числе с использованием внешних инструментов.

Таким образом, систему типов можно использовать в том числе как интерфейс взаимодействия внешних инструментов статического анализа с интегрированной средой разработки.

В дальнейшем для простоты изложения любые данные, записываемые в репозиторий по некоторому аспекту, будем называть типами по этому аспекту.

3.2.2 Многократная запись в репозиторий

При объявлении аспекта он может быть определён либо как аспект с однократной записью в репозиторий, либо как аспект, поддерживающий многократную запись.

Аспекты с однократной записью запрещают запись в репозиторий для вершины, которой уже приписан тип по данному аспекту. Если же тип по данному аспекту вершине не приписан, то запрещена операция чтения. Таким образом, текущий тип вершины по данному аспекту может находиться в двух состояниях: ещё не записан или уже записан. При этом, если система типов детерминирована, то операция чтения типа вершины по данному значению, если она возможна, вернёт всегда одно и то же значение.

Аспекты, разрешающие многократную запись, не накладывают подобных ограничений на операции записи и чтения — запись и чтение типа по этому аспекту всегда считается допустимым. В таком случае возникает два вопроса.

1. Что вернёт операция чтения, если до этого записи не производилось?
2. Что останется в репозитории после операции записи, если в репозитории уже был записан тип для этой вершины?

Ответ на эти вопросы можно предоставить, если при объявлении аспекта, разрешающего многократную запись, указать моноид, содержащий значение типа по этому аспекту перед первой записью и ассоциативную функцию, вычисляющую новое значение на основе значения, которое было в репозитории до этого, и значения, переданного в качестве параметра в функцию записи.

Например, для системы типов Хиндли-Милнера единицей моноида является свободная переменная типа, а операцией умножения — структурная унификация двух типовых деревьев.

#Пример: Конъюнкция требований#

3.3 Правила типизации

Основу описания систем типов на разработанном языке составляют локальные правила типизации. Правило типизации состоит из *шаблона*

AST, по которому определяется применимость правила, специальных *аннотаций* и фрагмента программного кода, непосредственно осуществляющего необходимые для вывода и проверки типов действия, — *действующей процедуры правила*.

Каждое правило применяется к тем фрагментам *AST*, которые соответствует шаблону этого правила. Если правило может быть применено к некоторому фрагменту (поддереву) *AST*, создаётся экземпляр этого правила, привязанный к корневой вершине этого поддерева, которая становится *корневой вершиной экземпляра*. Затем у созданного экземпляра правила вызывается действующая процедура. Результатом работы этой процедуры, а значит и результатом применения правила к фрагменту *AST*, является внесение изменений в репозиторий типов в отношении вершин из этого фрагмента.

Шаблоном может служить, к примеру, конкретная конструкция языка. Например, в большинстве императивных языков будет правило, применимое к операциям присваивания. Для каждого присваивания в типизируемой программе будет создан экземпляр этого правила. Действующая процедура правила прочитает из репозитория тип переменной из левой части присваивания, прочитает тип выражения из правой части и сравнит эти типы на совместимость.

Как видно из примера, правило может обращаться к репозиторию типов с запросами на чтение или запись. Но если запросы производятся по аспекту с одноразовой записью, то перед тем, как производить чтения типа некоторой вершины, этот тип должен быть записан в репозиторий в процессе выполнения другого правила. Это требование формирует достаточно строгие ограничения на порядок запуска экземпляров правил.

3.3.1 Аннотации правил

Требования к порядку применения правил содержатся в самих правилах в виде специального вида аннотаций. Минимальный набор аннотаций, позволяющий запускать правила в правильном порядке, состоит из одной функции, отвечающей на вопрос, может ли экземпляр правила быть выполнен в данный момент (достаточно ли уже вычислено информации для корректной работы данного экземпляра правила).

Но такой минимальный набор не всегда приводит к эффективной работе. Если появятся циклические зависимости, при которых ни одно правило не может быть выполнено, то это выяснится только в процессе выполнения правил. Также неэффективно после завершения работы каждого эк-

земпляра правил опрашивать все оставшиеся правила, чтобы выбрать из них то, которое должно быть выполнено следующим.

Чтобы определять порядок выполнения правил заранее, до начала их выполнения, используются дополнительные аннотации зависимостей правил. В начале, после создания всех необходимых экземпляров, каждый экземпляр предоставляет список вершин, тип которых должен быть в репозитории на момент запуска этого экземпляра. Также каждый экземпляр предоставляет список вершин, тип которых будет записан в репозиторий в результате работы этого экземпляра. На основе этих данных составляется направленный двудольный граф, вершинами которого являются, с одной стороны, вершины AST (по некоторому аспекту) и, с другой стороны, экземпляры правил. Дуга графа, идущая из экземпляра правила в вершину AST, означает, что это правило записывает (или, если дуга пунктирная, то предположительно записывает) в репозиторий тип этой вершины. Дуга графа, идущая из вершины AST в экземпляр правила, означает, что этому правилу требуется (или, если дуга пунктирная, то предположительно требуется) тип этой вершины.

Пример графа - рисунок

3.3.2 Подбор правил

В большинстве ситуаций не требуется вычислять и проверять типы для всего проекта. Например, компиляция крупных проектов происходит отдельными файлами или модулями. Но часто файл тоже является достаточно большой единицей типизации. Например, это может происходить при редактировании текста в IDE, когда при небольшом изменении кода в редакторе нет необходимости перевычислять типы всех выражений в файле. Также возможна ситуация, когда требуется вычислить тип одной конкретной вершины (например, эта необходимость возникает при генерации некоторых высокоуровневых языков). В этом случае нужно создавать и запускать только те экземпляры правил, которые участвуют в выводе типа нужной вершины.

Для решения задачи подбора правил используются обратные аннотации зависимостей правил. Обычных (прямых) аннотаций зависимостей правил для этой цели недостаточно, так как прямые аннотации зависимостей вычисляют по экземпляру правила вершины AST, используемые этим правилом, то есть строят двудольный граф зависимостей со стороны экземпляров правил. А в ситуации, когда экземпляры правил ещё не со-

зданы, и требуется найти те из них, которые нужно создать, то же граф должен строиться со стороны вершин AST.

3.3.3 Автоматическая генерация аннотаций

Локальные правила вывода применяются в определённом порядке, так, чтобы информация, необходимая для работы некоторого правила, была занесена в репозиторий типов правилами, применёнными до этого.

Для того, чтобы корректно определять порядок выполнения правил, необходим способ определения для каждого правила информации, необходимой для работы этого правила, и информации, записываемой в репозиторий в результате работы этого правила. Если правила описываются пользователем на языке общего назначения, то они могут быть снабжены соответствующими аннотациями, если же правила генерируются из предметно-ориентированного языка, то аннотации могут быть получены на основе статического анализа кода.

Метаинформация, прилагающаяся к правилу, частично может быть выведена из кода правила методом статического анализа, что представляет собой перспективу для дальнейших исследований.

3.4 Операции над типами

Некоторые языки программирования при описании их системы типов используют понятие отношений на типах. Наиболее распространённым отношением является отношение подтипа, однако в языке могут существовать и несколько отношений. Например, в языке Java 2 используются 8 отношений. В проектируемой системе предлагается описывать отношения с помощью более общей концепции операций с типами. В зависимости от целей использования бинарное отношение можно задать в виде бинарной булевой функции или в виде функции от первого аргумента, возвращающей последовательность допустимых вторых аргументов.

Так как система проектируется так, чтобы допустить возможность расширения набора используемых типов, то отношения также должны объявляться расширяемо. Для этого предложено разделять объявление операций над типами от объявлений реализации этих операций.

Объявление операции на типах содержит количество аргументов этой операции и тип результата.

Реализаций одной операции может быть несколько для разных шаблонов аргументов.

Реализация n -арной операции содержит n шаблонов аргументов и функцию, непосредственно производящую вычисление.

Если реализации не пересекаются по шаблонам аргументов, то операция является функцией (возможно, частично определённой). Если же к некоторым аргументам могут быть применимы одновременно несколько реализаций, то возникает вопрос о том, которая из них должна быть применена.

В противном случае возможны 2 разные политики вычисления в зависимости от вычисляемой операции. Политика приоритетов предполагает некорректное завершение вычисления, если нет ни одной подходящей реализации. В случае же нескольких подходящих реализаций вызывается та из них, которая имеет наибольший приоритет.

Альтернативой политики приоритетов является политика равноправных реализаций. Операция определяет коммутативный моноид, содержащий значение по умолчанию, возвращаемое для тех аргументов, для которых не нашлось ни одной подходящей реализации, и бинарную функцию, по которой производится свёртка результатов работы всех подходящих реализаций, если подходящих реализаций несколько.

3.5 Типовые переменные

Типовые переменные могут использоваться для двух целей [tapl, 22.2]: представление полиморфных типов и использование в промежуточных вычислениях во время вывода типов.

В разработанном языке описания систем типов при создании переменной не требуется указывать, с какой из вышеуказанных целей будет использоваться переменная (будет ли её значение вычислено, или над ней появится квантор всеобщности). После создания переменной она является свободной, то есть её значение не определено. В дальнейшем её значение может стать конкретным типом или типом, выраженным через другие переменные. Это может произойти через операцию присваивания значения типовой переменной. После того, как переменной присвоено значение, она становится зависимой переменной, и в каждом месте, где она используется, она заменяется на своё значение. На практике разумно заменять упоминания переменной на её значение не непосредственно в тот момент, когда переменная перестаёт быть свободной, а лениво, при первом запросе

к её упоминанию. В результате такой замены все переменные, упоминания которых можно наблюдать в репозитории, являются свободными.

Если же значение переменной так и не будет определено в процессе вывод типов, то есть после выполнения всех экземпляров правил вывода, которые могли присвоить её значение, она всё равно осталась свободной, значит она может принимать любые значения. Как только об этом становится известно, переменная перемещается в список переменных некоторого полиморфного (содержащего переменную) типа.

Принцип локальности подсказывает, что переменные типов чаще используются вблизи того места в дереве, где они были объявлены. Поэтому логично, создавая переменную, привязывать её к определённом месту в дереве. А так как, если переменной не был присвоен тип, она становится частью полиморфного типа некоторой вершины, то логично привязывать её к той вершине, частью полиморфного типа которой она может стать.

Глава 4

Правила вывода типов для языка Haskell

Каждое правило вывода типов объявляется для одной конструкции языка. В процессе вывода к каждой вершине этой конструкции будет применён соответствующий экземпляр правила. Рассмотрим основные конструкции упрощённого языка Haskell и правила для них.

4.1 Объявление алгебраических типов

#схема# Здесь нужно выставить типы конструкторам. Конструктор с n аргументами будет иметь функциональный тип с n аргументами, причём аргументы и результат копируются непосредственно из объявления.

Замечание 4.1.1. Так как алгебраические типы могут иметь параметры, то конструкторы могут иметь полиморфный тип атваризатрав. Для удобного представления потенциально полиморфных типов в виде AST создан концепт `PolymorphicType`, состоящий из дерева обычного типа (возможно, содержащего переменные) и списка использованных переменных. Введение особой конструкции для потенциально полиморфных типов является лишь удобным способом организовать привязку (binding) переменных типа.

Замечание 4.1.2. В текущей версии типизируемого языка не поддерживаются классы типов. Все переменные типов не имеют ограничений.

Замечание 4.1.3. Проверка type kinds не влияет на вывод типов в корректно составленной программе. Поэтому не будем включать эту проверку в правила вывода типов. Для проверок, от которых не зависит основной процесс вывода типов, будем создавать отдельные правила, чтобы обеспечить модульность системы типов (например, проверки можно в некоторых ситуациях отключать для ускорения вычисления типов, если требуется не найти все ошибки типов, а вычислить конкретный тип. Проверка type kinds потребует добавить дополнительные поля в концепт переменной типа, это будет впоследствии уточнено.

4.2 Лямбда-выражение

#схема#: лямбда-выражение lambda, его аргумент arg, тело функции body. Правило на псевдокоде:

```
argType = new fresh type variable; bodyType = new fresh type variable;
typeof(arg) := argType; typeof(body) := bodyType; typeof(lambda) := "argType -> bodyType";
```

Замечание 4.2.4. Присваивание типов и resolve-процедура. Наиболее важной частью правил являются присваивания типов. При этом правило может присваивать тип не только той вершине, для которой создан конкретный экземпляр правила, но и вершинам в её поддереве, к которым заранее указан путь из этой вершины (запретить присваивание типов произвольно найденным вершинам предложено в [Elovkov D.]). В результате одной и той же вершине тип может присваиваться несколько раз различными правилами. Если это происходит, запускается resolve-процедура, вычисляющая конъюнкцию требований к типу.

Resolve-процедура принимает на вход вершину и два типа: старый, который уже был присвоен вершине, и новый, попытка присваивания которого вершине привела к вызову resolve-процедуры. В общем случае resolve-процедура должна задаваться в виде набора правил, но для данной конкретной задачи её можно задать одновременно, так как здесь resolve-процедура представляет собой процедуру унификации. Унификация двух алгебраических типов для различных типов возвращает ошибку, а для одинаковых запускает унификацию для соответствующих параметров. Аналогично для функциональных типов и применений типов. Если вызвана унификация свободной переменной и не переменной (алгебраического типа данных, функционального типа, применения типов), то пере-

менная перестаёт быть свободной и ей присваивается конкретное значение. Если же вызвана унификация двух свободных переменных, то они обе выражаются через новую переменную. Создание новой переменной вместо выражения одной переменной через другую преследует несколько целей. Во-первых, при дальнейшем добавлении классов типов новая переменная будет совмещать в себе классы обеих исходных переменных. Во-вторых, новая переменная создаётся привязанной к минимальной общей надвершине вершин, к которым привязаны исходные переменные. #схема, пример с мономорфизмом#

4.3 Правило для всех выражений

Для каждого выражения $expr$ применим правило: $aType = \text{new fresh type variable}; \text{typeof}(expr) := aType$;

Это правило позволяет правильно вывести тип тогда, когда других правил нет. Например, для объявления $undefined = undefined$. В остальных случаях это правило не влияет принципиально на вывод типов.

Замечание 4.3.1. Type kinds и resolve-процедура. Вышеописанное правило может быть полезно при добавлении проверок type kinds. Так как известно, что тип любого выражения должен иметь kind $*$, то это можно задать непосредственно при создании переменной. Дополнительно потребуется ввести проверку совпадения type kinds у аргументов resolve-процедуры.

4.4 Применение функции к аргументу

#схема#: выражение применения функции к аргументу app , функция fun , аргумент arg . $argType = \text{new fresh type variable}; resType = \text{new fresh type variable}; \text{typeof}(arg) := argType; \text{typeof}(app) := resType; \text{typeof}(fun) := "argType \rightarrow resType"$;

4.5 Ссылка на переменную лямбда-выражения

#схема#: ссылка ref на переменную $decl$. Так как все упоминания переменной должны иметь одинаковый тип, логично приравнять тип ссыл-

ки типу объявления. $aType = \text{new fresh type variable}$; $\text{typeof}(\text{ref}) := aType$; $\text{typeof}(\text{decl}) := aType$;

4.6 Ссылка на объявление выражения или конструктора

Как было указано выше, объявления выражений или конструкторов (далее – объявления, *bindings*) имеют потенциально полиморфный тип, то есть содержат, помимо собственно типа, список переменных типа, которые могут принимать произвольное значение. При этом ссылка на объявление будет иметь такой же тип, но некоторые переменные в нём могут быть конкретизированы. Заметим, что в разных ссылках на одно и то же объявление переменные могут быть конкретизированы по-разному. Подобное поведение можно осуществить, копируя переменные из полиморфного типа в общее пространство переменных. Новые переменные создаются привязанными к ближайшей вершине с потенциально полиморфным типом.

4.7 Объявление выражения

В объявлении выражения производится обратная процедура: переменные копируются из общего пространства переменных в список переменных полиморфного типа. При этом захватываются только те переменные, которые привязаны к вершине не выше рассматриваемой.

Замечание 4.7.1. О порядке выполнения правил и окончательно определённых типах. Так как объявления выражений являются частью видимого извне интерфейса модуля, то запросы об их типе могут приходиться из различных источников. Поэтому тип объявлений должен быть определён корректно или ещё не определён – промежуточные вычисления не должны сказываться на этом типе. Для того, чтобы этого достичь, необходимо запускать копирование типов после того, как тип самого выражения окончательно определён. Окончательно определённым типом можно считать тогда, когда становится понятным, что все переменные, привязанные к рассматриваемой вершине или ниже, уже не будут конкретизированы. Фактически, окончательная определённость типа даёт возможность произвести захват всех свободных переменных, привязанных к поддереву, и поставить над ними квантор всеобщности. Конкретные алгоритмы, позволя-

ющие установить момент, когда тип окончательно определён, и запустить правило, копирующее переменные, требуют отдельного исследования графов зависимостей вершин и правил.

4.8 Выражение `let`

Свойство выражения `let`, которое ещё не учтено другими правилами, заключается в том, что тип выражения `let` совпадает с типом выражения, стоящего после `in`. $aType = \text{new fresh type variable}; \text{typeof}(\text{letExpr}) := aType; \text{typeof}(\text{innerExpr}) := aType;$

Замечание 4.8.1. О мономорфизме и окончательно определённых типах. Выражение `let` имеет одну примечательную особенность: внутри его объявлений могут находиться ссылки на переменные объёмлющих функций. В результате может оказаться, что тип некоторого объявления внутри `let` зависит от переменных, привязанных к вышележащим вершинам. В этом случае тип этого значения этого объявления, после того как он будет признан окончательно определённым, будет зависеть от этих переменных. Это не мешает скопировать переменные, привязанные к рассматриваемому объявлению, в полиморфный тип, который также будет зависеть от переменных, привязанных к вышележащим вершинам. В конечном итоге эти переменные могут быть конкретизированы, тогда полиморфизма по ним не будет. Или, если они не будут конкретизированы, они скопируются в некоторый вышележащий полиморфный тип.

Пример (классический, из Haskell wiki). $f\ x = \text{let } g\ y\ z = ([x,y],z) \text{ in } (g\ (), g\ \text{True}\ ())$ Данный пример корректно типизировать невозможно, так как тип функции `g` связан (зависит от общей переменной типа) с типом переменной `x`. Из использования `g` в выражении `g\ ()` следует, что переменная `x` имеет тип `()`, а из использования `g` в выражении `g\ True\ ()` следует, что переменная `x` имеет тип `Bool`. А разные алгебраические типы не совместимы друг с другом – `resolve`-процедура, пытающаяся провести унификацию, возвращает ошибку.

4.9 Выражение `case` и сопоставление с образцом

Для образцов правила типизации аналогичны правилам для выражений: ссылки на конструкторы копируют в локальный контекст тип кон-

строуктора, применение конструктора к аргументу (тоже образцу) аналогично применению функции к аргументу. Заметим, что количество аргументов у конструктора в образце должно совпадать с количеством аргументов в объявлении конструктора. И обеспечить это должна система типов (особенно, если иметь в виду расширение GADTs). Один из способов, которым это можно обеспечить, заключается в дополнительном присваивании образцу нужного алгебраического типа. В результате, если количество аргументов недостаточно, то этот тип будет конфликтовать в выведенным функциональным.

После того, как выведены типы образцов (окончательно определены), их можно скопировать на тип сравниваемого с образцами выражения. Если не сойдутся – ошибка.

Тип самого выражения `case` можно определить, приравняв его (присвоив одну и ту же переменную) к типам ветвей.

Глава 5

Описание системы типов языка Java

Заключение

Создан прототип движка, выполняющего правила в нужном порядке в соответствии с условиями, указанными в правилах. Реализован набор правил, осуществляющий вычисление и проверку типов для системы Хиндли-Милнера с *let*-полиморфизмом (упрощённая версия языка Haskell). При этом расширение Haskell, вводящее обобщённые алгебраические типы данных (GADTs) сводится к добавлению к базовому набору двух новых правил. Тип *Integer* и другие предопределённые в языке типы также вводятся в виде расширений языка, добавляющих новые правила системы типов.

Нетривиальной задачей является моделирование декларативной по своей природе системы Хиндли-Милнера в виде императивных правил. Однако следует заметить, что системе Хиндли-Милнера могут соответствовать различные алгоритмы вычисления типов, например, алгоритм *W* и алгоритм *M* [3]. При корректной программе эти алгоритмы приведут к одинаковому результату, но при наличии ошибок сообщения об ошибках будут разными. Различие между алгоритмами *W* и *M* заключается в порядке решения уравнений, и может быть выражено именно при императивном подходе.

Также создан набор правил, осуществляющий вычисление типов для упрощённой версии языка Java (без параметрического полиморфизма и перегрузки методов).

Основным критерием качества разрабатываемой системы описания правил является возможность описывать правила в тех же терминах, которые использовались при разработке языка. Для языка Haskell такими терминами являются уравнения над типами, а для языка Java – отношения и операции над типами. Соответственно, язык описания правил можно считать пригодным для описания систем типов Haskell и Java, если в

правилах для языка Haskell правила будут состоять в основном из кода, воспринимаемого как уравнения, а правила для Java будут оперировать отношениями и операциями над типами.

Разработана система языковых и библиотечных конструкций, позволяющая записывать алгоритмы вычисления и проверки типов в виде правил, структурно повторяющих пункты документации, по крайней мере для двух различных языков программирования. В качестве иллюстрации реализованы упрощённые версии алгоритмов вычисления и проверки типов языков Haskell и Java. Разработанная система после некоторой технической доработки может быть использована для описания языков в платформах разработки языков, таких как JetBrains MPS.