

Федеральное государственное автономное образовательное учреждение
высшего образования «Национальный исследовательский университет
«Высшая школа экономики»

Факультет Санкт-Петербургская школа физико-математических и
компьютерных наук

Департамент информатики

Основная образовательная программа

Прикладная математика и информатика

Кайсин Илья Сергеевич

Задача обитаемости в системах типов низкого ранга

Выпускная квалификационная работа

Допущена к защите.

Зав. кафедрой:

д. ф.-м. н., профессор Омельченко А. В.

Научный руководитель:

к. ф.-м. н., профессор Москвин Д. Н.

Рецензент:

Пеленицын А. М

Санкт-Петербург
2019

Оглавление

Аннотация	3
Введение	5
1. Система типов $\lambda_{\wedge\eta}$	6
1.1. Подтипизация	6
1.2. Существенность η правила	10
1.3. Выводы и результаты по главе	11
2. Населяющий алгоритм	12
2.1. Населяющий алгоритм для λ_{\wedge}	12
2.2. Населяющий алгоритм для $\lambda_{\wedge\eta}$	13
2.3. Отличия алгоритмов	13
2.4. Выводы и результаты по главе	13
3. Свойства алгоритма	15
3.1. Корректность	15
3.2. Полнота	15
3.3. Завершаемость	18
3.4. Выводы и результаты по главе	19
Список литературы	20

Аннотация

Типы являются важной абстракцией в языках программирования. Ограничивая множество корректных программ, они позволяют программистам писать более безопасный код и избегать ошибок. Теория типов – математический аппарат, формализующий данную область. Одним из базовых вопросов теории типов является *задача обитаемости*: «Существует ли терм (выражение в языке) заданного типа?». Задача обитаемости может быть поставлена в различных *системах типов* (моделях, описывающих часть языка программирования, относящуюся к типам). Цель данной работы – решить проблему обитаемости для некоторой определённой системы типов, а именно для просто типизированного лямбда исчисления с пересечениями второго ранга и подтипизацией. Недавние работы показывают, что задача обитаемости неразрешима в системах без ограничения на ранг [12] и EXPSPACE-полной в системе с пересечениями второго ранга, но без подтипизации [7, 13]. Однако система и с пересечениями второго ранга, и с подтипизацией до сих пор не была изучена. В данной работе мы показываем, что задача обитаемости разрешима в этой системе. Для этого мы приводим алгоритм, решающий её, доказывая его корректность, полноту и завершаемость. Помимо этого, работа закрывает некоторые неточности предыдущих статей и предлагает новые подходы к реализации описанных в них алгоритмов.

Ключевые слова: теория типов, лямбда исчисление, типы-пересечения, ранг типа, подтипизация, проблема обитаемости.

Types are an important abstraction in programming languages. They enable to program in a safer way and avoid bugs setting restriction on programs and distinguishing correct and incorrect ones. Type theory is the mathematical apparatus for this domain, one of whose basic questions is the inhabitation, usually defined as follows: «does there exist a term (some expression in the language) of a given type?». The question could be raised in different type systems (models of type-related programming languages parts), the goal of this project is to solve this question in some specific systems. Namely, simply typed lambda calculus with type intersection of rank two (type system restriction) and subtyping. Recently presented works show that the problem is undecidable in type systems without rank limitation [12] and EXPSPACE-complete without subtyping [7, 13], but the system with both of them is not studied yet. The approaches include an adaptation of an existing in similar type systems inhabitation algorithm, proof of its correctness and soundness, implementing and testing the algorithm in Haskell programming language. The preliminary results also cover several flaws in the previous works and suggest that the inhabitation algorithm scheme could be improved in terms of efficiency.

Keywords: type theory, lambda calculus, intersection types, type rank, subtyping, type inhabitation problem.

Введение

Почти во всех языках программирования в том или ином виде представлены типы. Чем сильнее система типов языка, тем больше ограничений она накладывает на корректные программы и тем больше ошибок может быть найдено на *этапе компиляции*, что позволяет уменьшить число потенциальных ошибок на *этапе исполнения*, когда программа работает в реальной среде, и цена каждой ошибки велика. Изучением систем типов и их свойств занимается теория типов.

Помимо языков программирования, теория типов применяется в системах интерактивных доказательств – средствах, позволяющих математикам полуавтоматически формулировать и доказывать теоремы на некотором формальном языке; используется для описания оснований математики как альтернатива наивной теории множеств.

Наиболее развитую систему типов имеют функциональные языки программирования. Основанием этой системы является типизированное *лямбда исчисление*.

Лямбда исчисление было впервые введено Алонзо Чёрчем в 1930-х годах и на данный момент имеет множество модификаций. Самая базовое из типизированных лямбда исчислений – просто типизированное лямбда исчисление. Типы в этой системе конструируются из двух составных частей: типовых переменных (например, `Int` – тип целых чисел или `String` – тип строк) и операции (типового конструктора) « \rightarrow ». « \rightarrow » представляет функциональный тип. Например, тип «`Int \rightarrow String`» – это тип функций, отображающих целые числа в строки.

Система типов, рассматриваемая в данной работе, содержит ещё один типовой конструктор – « \wedge », представляющий пересечение двух типов. Лямбда исчисление с пересечениями известно с 1960-х, значимость этой системы подтверждается многочисленными работами, посвященными ей: модели лямбда исчисления [1, 3], проблемы нормализации и оптимальной редукции [9, 10], вывод типов и компиляция [6, 14].

Другим важным аспектом систем типов с пересечениями является их логическая интерпретация. Изоморфизм Карри-Говарда ставит в соответствие типам логические высказывания, а термам – доказательства этих высказываний. Благодаря такой интерпретации, можно считать, что решая задачу обитаемости, мы по высказыванию изучаем его доказательства – ищем их или проверяем их наличие. Просто типизированное лямбда исчисление с пересечениями при этом соответствует пропозициональной логике с пересечениями. Неразрешимость задачи обитаемости в этой системе [11] говорит о её большой выразительной силе. Поэтому, в частности, важно разграничить случаи разрешимости и неразрешимости. В данной работе мы рассматриваем просто типизированное лямбда исчисление с пересечениями и подтипизацией. При этом выясняется, что, ограничивая определённым образом множество допустимых типов системы, мы ослабляем её, и задача обитаемости становится разрешимой.

1. Система типов $\lambda_{\wedge\eta}$

Система типов $\lambda_{\wedge\eta}$ отличается от просто типизированного лямбда исчисления новым типовым конструктором: \wedge , соответствующим пересечению типов. Эту операцию можно понимать в теоретико-множественном смысле: типом $\sigma \wedge \tau$ типизируются такие и только такие термы, которые типизируются и σ , и τ . Правила вывода, соответствующие этому поведению, обозначаются $(I\wedge)$ и $(E\wedge)$ (введение пересечения и элиминация пересечения соответственно).

Кроме того, вводится ещё одно дополнительное правило (η) , позволяющее проводить эта-экспансию.

Таким образом, система состоит из следующих правил вывода: Ax , $I \rightarrow$, $E \rightarrow$, $I\wedge$, $E\wedge$, η .

$$\begin{aligned} & \frac{}{\Gamma, x: \tau \vdash x: \tau} (Ax) \\ & \frac{\Gamma, x: \sigma \vdash M: \tau}{\Gamma \vdash (\lambda x.M): \sigma \rightarrow \tau} (I \rightarrow) \\ & \frac{\Gamma \vdash M: \sigma \rightarrow \tau \quad \Gamma \vdash N: \sigma}{\Gamma \vdash (MN): \tau} (E \rightarrow) \\ & \frac{\Gamma \vdash M: \sigma \quad \Gamma \vdash M: \tau}{\Gamma \vdash M: \sigma \wedge \tau} (I\wedge) \\ & \frac{\Gamma \vdash M: \sigma \wedge \tau}{\Gamma \vdash M: \sigma} (E\wedge) \\ & \frac{\Gamma \vdash M: \sigma}{\Gamma \vdash (\lambda x.Mx): \sigma} (\eta) \end{aligned} \tag{1}$$

1.1. Подтипизация

Определим на типах отношение подтипизации. С теоретико-множественной точки зрения подтипизация соответствует отношению «быть подмножеством»: если терм типизируется σ , то он типизируется всеми надтипами σ , то есть такими τ , что $\sigma \leq \tau$. Отношение определим следующими аксиомами и правилами, аналогично определе-

нию из [4].

$$\begin{aligned}
& \overline{\sigma \leq \sigma} \quad (A1) \\
& \overline{\sigma \leq \sigma \wedge \sigma} \quad (A2) \\
& \overline{\sigma \wedge \tau \leq \sigma} \quad (A3) \\
& \overline{\sigma \wedge \tau \leq \tau} \quad (A4) \\
& \overline{(\sigma \rightarrow \tau_1) \wedge (\sigma \rightarrow \tau_2) \leq \sigma \rightarrow (\tau_1 \wedge \tau_2)} \quad (A5) \\
& \frac{\sigma \leq \sigma' \quad \tau \leq \tau'}{\sigma \wedge \tau \leq \sigma' \wedge \tau'} \quad (R1) \\
& \frac{\sigma \leq \sigma' \quad \tau \leq \tau'}{\sigma' \rightarrow \tau \leq \sigma \rightarrow \tau'} \quad (R2) \\
& \frac{\tau_1 \leq \tau_2 \quad \tau_2 \leq \tau_3}{\tau_1 \leq \tau_3} \quad (R3)
\end{aligned} \tag{2}$$

В [5] показано, что правило (η) может быть заменено следующим правилом:

$$\frac{\Gamma \vdash M : \sigma \quad \sigma \leq \tau}{\Gamma \vdash M : \tau} (\leq)$$

На самом деле, поскольку $\sigma \wedge \tau \leq \sigma$ и $\sigma \wedge \tau \leq \tau$, правило $(E\wedge)$ также избыточно. Таким образом, правила в этой системе следующие:

$$\begin{aligned}
& \overline{\Gamma, x : \tau \vdash x : \tau} \quad (Ax) \\
& \frac{\Gamma, x : \sigma \vdash M : \tau}{\Gamma \vdash (\lambda x.M) : \sigma \rightarrow \tau} \quad (I \rightarrow) \\
& \frac{\Gamma \vdash M : \sigma \rightarrow \tau \quad \Gamma \vdash N : \sigma}{\Gamma \vdash (MN) : \tau} \quad (E \rightarrow) \\
& \frac{\Gamma \vdash M : \sigma \quad \Gamma \vdash M : \tau}{\Gamma \vdash M : \sigma \wedge \tau} \quad (I\wedge) \\
& \frac{\Gamma \vdash M : \sigma \quad \sigma \leq \tau}{\Gamma \vdash M : \tau} (\leq)
\end{aligned} \tag{3}$$

Полученную систему будем называть λ_{\leq} . Она эквивалентна $\lambda_{\wedge\eta}$ в смысле типизации: утверждение типизации, верное в системе λ_{\leq} , верно в $\lambda_{\wedge\eta}$ и наоборот. Поэтому в контексте нашей задачи эти две системы полностью взаимозаменяемы.

Определим отношение эквивалентности на типах следующим образом.

Определение 1. $\sigma \sim \tau \iff \sigma \leq \tau \text{ и } \tau \leq \sigma$

Благодаря правилу (\leq) , для эквивалентных типов верно следующее утверждение:

Лемма 1.1. *Множества термов, типизируемых эквивалентными типами, в точности совпадают.*

В частности, это означает, что переход к эквивалентному типу не влияет на его обитаемость.

Легко видеть, что отношение « \sim » является отношением конгруэнтности (в смысле [2]) а именно, верна следующая Лемма:

Лемма 1.2. Пусть $\alpha, \beta, \gamma, \delta$ – типы. Если $\alpha \sim \beta$ и $\gamma \sim \delta$, то $(\alpha \rightarrow \gamma) \sim (\beta \rightarrow \delta)$, а также $(\alpha \wedge \gamma) \sim (\beta \wedge \delta)$.

Для удобства введём следующее обозначение:

Определение 2.

$$\hat{\tau} = \{\tau_i \mid \tau_1 \wedge \dots \wedge \tau_k = \tau$$

и τ_i не является пересечением ни для какого $i\}$

То есть $\hat{\tau}$ – множество, полученное разделением всех пересечений на верхнем уровне τ .

В [4] показана следующая эквивалентность:

Лемма 1.3. $\alpha \rightarrow \beta \wedge \gamma \sim (\alpha \rightarrow \beta) \wedge (\alpha \rightarrow \gamma)$

Применяя её многократно, мы можем привести тип к нормальной форме. А именно, введём операцию « $*$ » следующим образом:

Определение 3.

$$\alpha^* = \alpha$$

если α – типовая переменная

$$(\sigma \wedge \tau)^* = \sigma^* \wedge \tau^*$$

$$(\sigma \rightarrow \tau)^* = \bigwedge_{\varphi \in \hat{\tau}^*} \sigma \rightarrow \varphi$$

Например, $(\alpha \rightarrow \gamma \wedge (\beta \rightarrow \gamma \wedge \delta))^* = (\alpha \rightarrow \beta \rightarrow \gamma) \wedge (\alpha \rightarrow \beta \rightarrow \delta) \wedge (\alpha \rightarrow \gamma)$.

Операция « $*$ » переводит тип в эквивалентный.

Лемма 1.4. Для любого типа τ : $\tau \sim \tau^*$

Определение 4. Тип τ находится в нормальной форме, если $\tau^* = \tau$.

Общий вид типа в нормальной форме после применения « $*$ » описывается следующей леммой:

Лемма 1.5. Для любого типа τ : $\tau^* = \bigwedge_i (\varphi_{i1} \rightarrow \dots \rightarrow \varphi_{iki} \rightarrow \alpha_i)$, где α_i – типовая переменная.

Следствие 1.5.1. *Типы в нормальной форме без пересечений на верхнем уровне — в точности типы вида $\dots \rightarrow \alpha$, где α — некоторая типовая переменная.*

В [4] показано, что отношение (\leq) рекурсивно и существует алгоритм, позволяющий определить для двух типов α и β верно ли, что $\alpha \leq \beta$:

Алгоритм 1.

$$\begin{aligned}
&\alpha \leq \beta = (\alpha == \beta) \\
&\text{если } \alpha \text{ и } \beta \text{ — типовые переменные} \\
&\sigma \leq \tau = \text{False} \\
&\text{если один из типов — переменная, а другой тип стрелочный} \\
&(\sigma_1 \rightarrow \tau_1) \leq (\sigma_2 \rightarrow \tau_2) = (\tau_1 \leq \tau_2) \text{ AND } (\sigma_2 \leq \sigma_1) \\
&\sigma \leq \tau = \forall \tau_i \in \hat{\tau}^*: \exists \sigma_j \in \hat{\sigma}^*: \sigma_j \leq \tau_i
\end{aligned}$$

Введём ещё несколько обозначений.

Определение 5. *Обозначим множество всех подтипов τ через $\underline{\tau}$, а множество всех его надтипов через $\bar{\tau}$. То есть $\underline{\tau} = \{\sigma \mid \sigma \leq \tau\}$, $\bar{\tau} = \{\sigma \mid \tau \leq \sigma\}$.*

Далее будем считать, что надтипы и подтипы могут использоваться везде, где могут использоваться типы. При этом операции над типами «поднимаются» на уровень множеств. Так, например, выражение $\alpha \rightarrow \bar{\beta} \rightarrow \underline{\gamma}$ следует понимать как множество $\{\alpha \rightarrow \beta' \rightarrow \gamma' \mid \beta' \in \bar{\beta}, \gamma' \in \underline{\gamma}\}$.

Утверждение о типизации, в котором фигурируют надтипы или подтипы, будем считать выводимым, если оно выводимо для некоторых элементов соответствующих множеств надтипов и подтипов.

Запись о вхождении типизации в контекст $(x : T) \in \Gamma$ в случае, если в T фигурируют надтипы или подтипы, означает, что $(x : \tau) \in \Gamma$ для некоторого $\tau \in T$.

Запись вида $T_1 \leq T_2$, где в T_i могут фигурировать подтипы и надтипы, означает, что $\forall \tau_1 \in T_1, \tau_2 \in T_2 : \tau_1 \leq \tau_2$.

Лемма 1.6. *Пусть $\sigma \rightarrow \tau$ — тип в нормальной форме. Тогда $\underline{\sigma \rightarrow \tau} = \varphi \wedge (\bar{\sigma} \rightarrow \underline{\tau})$, для некоторого (возможно, пустого) φ .*

Доказательство. Включение $\varphi \wedge (\bar{\sigma} \rightarrow \underline{\tau}) \subseteq \underline{\sigma \rightarrow \tau}$ почти очевидно:

$$\frac{\frac{\sigma \leq \bar{\sigma} \quad \underline{\tau} \leq \tau}{\bar{\sigma} \rightarrow \underline{\tau}} (R2)}{\varphi \wedge (\bar{\sigma} \rightarrow \underline{\tau}) \leq \sigma \rightarrow \tau} (A4)$$

Для включения $\underline{\sigma \rightarrow \tau} \subseteq \varphi \wedge (\bar{\sigma} \rightarrow \underline{\tau})$ достаточно посмотреть на Алгоритм 1. Подтип $\sigma \rightarrow \tau$ — это либо тип из $(\bar{\sigma} \rightarrow \underline{\tau})$ (третий случай алгоритма), либо тип-пересечение, один из элементов которого — подтип $\sigma \rightarrow \tau$; остальные элементы можно «переместить» в φ перестановкой. \square

1.2. Существенность η правила

Насколько добавление η - правило меняет систему? Система λ_{\wedge} существенно слабее $\lambda_{\wedge\eta}$, а именно, есть тип, обитаемый в системе без эта-правила и обитаемый в системе с ним.

Лемма 1.7. *Утверждение о типизации $x : \alpha \rightarrow \beta \wedge \gamma \vdash x : \alpha \rightarrow \beta$ верно в $\lambda_{\wedge\eta}$ но неверно в λ_{\wedge} .*

Доказательство. Докажем выводимость в системе $\lambda_{\wedge\leq}$:

$$\frac{x : \alpha \rightarrow \beta \wedge \gamma \vdash x : \alpha \rightarrow \beta \wedge \gamma \quad \alpha \rightarrow \beta \wedge \gamma \leq \alpha \rightarrow \beta}{x : \alpha \rightarrow \beta \wedge \gamma \vdash x : \alpha \rightarrow \beta}$$

Чтобы доказать, что утверждение неверно в системе λ_{\wedge} , предположим обратное. Рассмотрим последнее правило вывода, которое могло быть применено, чтобы получить данное утверждение. $(I \rightarrow)$ и $(E \rightarrow)$ не могли быть применены, поскольку они типизируют абстракцию и аппликацию соответственно, а x — типовая переменная. Правило (Ax) требует наличия соответствующей типизации в контексте, правило $(I\wedge)$ приписывает терму тип-пересечение, а не стрелочный тип.

Таким образом, единственным возможным правилом могло быть $(E\wedge)$:

$$\frac{\begin{array}{c} \vdots \\ x : \alpha \rightarrow \beta \wedge \gamma \vdash x : (\alpha \rightarrow \beta) \wedge \sigma \end{array}}{x : \alpha \rightarrow \beta \wedge \gamma \vdash x : \alpha \rightarrow \beta}$$

Какие правила вывода могли привести к данному утверждению о типизации? Только $(I \rightarrow)$ и $(E \rightarrow)$. Однако легко видеть, что их «обратное» применение порождает утверждение вида $x : \alpha \rightarrow \beta \wedge \gamma \vdash x : (\alpha \rightarrow \beta) \wedge \sigma$ для некоторого (возможно, пустого) σ , но утверждение такого вида уже встречалось ранее, и могло быть получено лишь с помощью $(I \rightarrow)$ и $(E \rightarrow)$. Значит, дерева вывода не существует. □

Лемма 1.8. *Тип $\delta \wedge (\alpha \rightarrow \beta \wedge \gamma) \rightarrow \delta \wedge (\alpha \rightarrow \beta)$ пуст в λ_{\wedge} , но содержит терм id в $\lambda_{\wedge\eta}$.*

Доказательство. Анализом возможных деревьев вывода, аналогичным рассуждениям в предыдущей лемме, получаем, что для существования типизации $\vdash M : \delta \wedge (\alpha \rightarrow \beta \wedge \gamma) \rightarrow \delta \wedge (\alpha \rightarrow \beta)$ необходима и достаточна типизация $x : \alpha \rightarrow \beta \wedge \gamma \vdash x : \alpha \rightarrow \beta$. □

Замечание 1.8.1. *Тип $(\alpha \rightarrow \beta \wedge \gamma) \rightarrow \alpha \rightarrow \beta$ обитает в обеих системах. В λ_{\wedge} он содержит $\lambda a.b.a$, но не $\lambda a.a$.*

1.3. Выводы и результаты по главе

В данной главе была рассмотрена система типов $\lambda_{\wedge\eta}$ и эквивалентная ей $\lambda_{\wedge\leq}$, для которых в следующих разделах будет приведён и доказан населяющий алгоритм. Также было введено отношение подтипизации « \leq »; отношение эквивалентности « \sim »; операции $\bar{\tau}$ и $\underline{\tau}$, ставящие в соответствие типу множество его надтипов и подтипов соответственно; операция « $*$ », приводящая тип к нормальной форме. Алгоритм приведения к нормальной форме является важной составной частью населяющего алгоритма, описанию которого посвящена следующая глава.

2. Населяющий алгоритм

В [7] был представлен населяющий алгоритм для системы λ_{\wedge} . Однако этот алгоритм описан недостаточно полно, что выяснилось при его реализации. Кроме того, алгоритм и доказательство его корректности содержат ошибку. Данная работа устраняет эти неточности и ошибки. Алгоритм для $\lambda_{\wedge\leq}$ является модификацией алгоритма из [7], поэтому сначала рассмотрим его (в исправленном варианте).

2.1. Населяющий алгоритм для λ_{\wedge}

Алгоритм 2. В процессе алгоритма решается система из нескольких задач: $[\Gamma_1 \vdash X : \tau_1, \dots, \Gamma_n \vdash X : \tau_n]$. Решением системы является терм X , удовлетворяющий всем утверждениям типизации одновременно. При этом поддерживается инвариант: все задачи системы имеют общий набор переменных в контексте (при этом одной и той же типовой переменной могут соответствовать разные типы в разных контекстах). Алгоритм заключается в применении следующих преобразований до тех пор, пока это возможно:

- 1. Один из типов τ_i – пересечение ($\tau_i = \sigma \wedge \rho$). Тогда i -я задача $\Gamma_i \vdash X : \tau_i$ заменяется двумя: $\Gamma_i \vdash X : \sigma$ и $\Gamma_i \vdash X : \rho$. Таким образом, размер системы увеличивается
- 2. Все типы τ_i стрелочные ($\tau_i = \sigma_i \rightarrow \rho_i$, $i = 1 \dots n$). Тогда решение системы – $X = \lambda x. X'$, где X' – решение новой системы $[(\Gamma_1, x : \sigma_1 \vdash X' : \rho_1), \dots, (\Gamma_n, x : \sigma_n \vdash X' : \rho_n)]$. То есть в этом случае во всех типах редуцируется первый аргумент.
- 3 $_{\wedge}$. Один из типов τ_i – это типовая переменная. Тогда X не может быть абстракцией и должен быть (возможно пустой) аппликацией некоторой головной переменной, взятой из контекста, к термам. В этом случае необходимо недетерминированно выбрать из контекста головную переменную x и число $k \geq 0$ таким образом, чтобы $\Gamma_i \vdash \lambda z_1 \dots z_k. x z_1 \dots z_k : \rho_i^1 \rightarrow \dots \rightarrow \rho_i^k \rightarrow \tau_i$ для всех $i = 1 \dots n$. Тогда решение – $X = x Z^1 \dots Z^k$, где Z^i – решение системы $[(\Gamma_1 \vdash Z^i : \rho_1^i), \dots, (\Gamma_n \vdash Z^i : \rho_n^i)]$. Здесь k систем независимы и могут быть решены параллельно.

Если ни одно из преобразований применить невозможно, то система не имеет решения. «Точка выхода» из алгоритма – тот случай в преобразовании 3, при котором $k = 0$, при этом решением является переменная.

В данном алгоритме не совсем ясным является преобразование 3 $_{\wedge}$: как именно выбираются ρ_i^j . Типизация $\Gamma_i \vdash \lambda z_1 \dots z_k. x z_1 \dots z_k : \rho_i^1 \rightarrow \dots \rightarrow \rho_i^k \rightarrow \tau_i$ не гарантирует того, что тип x в контексте Γ_i обязан быть $\rho_i^1 \rightarrow \dots \rightarrow \rho_i^k \rightarrow \tau_i$.

2.2. Населяющий алгоритм для $\lambda_{\wedge\eta}$

Алгоритм для системы с η -правилом во многом повторяет алгоритм для λ_{\wedge} . Принципиальное отличие заключается в преобразовании 3.

Алгоритм 3. *Поддерживается система из нескольких задач: $\Gamma_1 \vdash X : \tau_1, \dots, \Gamma_n \vdash X : \tau_n$. До тех пор, пока это возможно, выполняется одно из четырёх преобразований:*

- 0. Один из типов τ_i не находится в нормальной форме. Тогда заменим τ_i на τ_i^* .
- Шаги 1 и 2 аналогичны Алгоритму 2.
- 3 $_{\wedge\eta}$. Один из типов τ_i – это типовая переменная. Тогда X не может быть абстракцией и должен быть (возможно, пустой) аппликацией некоторой головной переменной, взятой из контекста, к термам. В этом случае необходимо недетерминированно выбрать переменную x и число $k \geq 0$ таким образом, чтобы $\Gamma_i \vdash x : \rho_i^1 \rightarrow \dots \rightarrow \rho_i^k \rightarrow \tau_i$ для всех $i = 1 \dots n$. Тогда решение – $X = xZ^1 \dots Z^k$, где Z^i – решение системы $(\Gamma_1 \vdash Z^i : \rho_1^i), \dots, (\Gamma_n \vdash Z^i : \rho_n^i)$. Здесь k систем независимы и могут быть решены параллельно.

Следствие 3.1.3 описывает, как именно устроен недетерминированный выбор x , k и ρ_i^j .

Аналогично алгоритму 2, если ни одно из преобразований применить невозможно, система не имеет решения.

2.3. Отличия алгоритмов

В [7] в качестве населяющего алгоритма для системы λ_{\wedge} приведён алгоритм, почти совпадающий с Алгоритмом 3, описанным выше. Однако на следующем примере можно убедиться, что данный алгоритм не является полным в системе λ_{\wedge} .

Пусть задача – $p : \alpha \rightarrow \beta \wedge \gamma, q : \alpha \vdash X : \beta$. Алгоритм должен произвести преобразование 3 $_{\wedge\eta}$, но в контексте нет такого x , что $p : \alpha \rightarrow \beta \wedge \gamma, q : \alpha \vdash x : \dots \rightarrow \beta$, поскольку, согласно Лемме 1.7, утверждение о типизации $p : \alpha \rightarrow \beta \wedge \gamma \vdash p : \alpha \rightarrow \beta$ неверно в λ_{\wedge} . Поэтому алгоритм завершится, не найдя решение $X = pq$. Но преобразование 3 $_{\wedge}$ применить можно: $p : \alpha \rightarrow \beta \wedge \gamma, q : \alpha \vdash \lambda z.pz : \alpha \rightarrow \beta$. Здесь λ позволяет применить $(I \rightarrow)$ и удалить \rightarrow .

2.4. Выводы и результаты по главе

В данной главе был рассмотрен населяющий Алгоритм 2 для системы λ_{\wedge} , описанный ранее в [7], а также приведён оригинальный населяющий Алгоритм 3 для

системы $\lambda_{\wedge\eta}$. Ключевым отличием алгоритмов является преобразование 3, некоторые подробности которого для Алгоритма 3 описывает следствие 3.1.3.

Довольно очевидным является тот факт, что результат, выдаваемый алгоритмами корректен: если на входе τ алгоритм выдаёт терм x , то $\emptyset \vdash x : \tau$. Однако кроме корректности нас интересует полнота (если тип обитаем, то алгоритм находит его обитателей) и завершаемость (если тип необитаем, алгоритм завершится за конечное число операций). Оказывается, завершаемость в таком виде доказать невозможно, поскольку в общем случае задача населённости для этой системы типов неразрешима [11]. Но, при условии некоторых ограничений на входные типы, алгоритм всё же гарантированно останавливается за конечное время. Описанию этих ограничений, а также доказательству корректности, полноты и завершаемости посвящена следующая глава.

3. Свойства алгоритма

В этом разделе будут доказаны свойства Алгоритма 3 для системы $\lambda_{\wedge\eta}$. А именно, его корректность (soundness), полнота (completeness), завершаемость (termination) на типах ранга ≤ 2 .

3.1. Корректность

Корректность алгоритма означает, что ответ, данный им, удовлетворяет входным условиям. То есть если некий терм был найден, то он действительно типизируется заданным типом.

Теорема 1 (Soundness). *Если алгоритм находит терм M для входного типа τ , то $\vdash M : \tau$*

Доказательство. Усилим утверждение: докажем, что алгоритм корректно решает систему задач. Доказательство – индукция по количеству шагов алгоритма. В базе индукции используется аксиома (Ax) . В переходах индукции в шагах 1, 2 и $3_{\wedge\eta}$ используются правила вывода $(I\wedge)$, $(I\rightarrow)$ и $(E\rightarrow)$ соответственно. \square

3.2. Полнота

Полнота – гораздо более содержательное свойство. Оно означает, что если существует терм заданного типа, то алгоритм обязательно найдёт или его, или другой терм этого типа. В нашем случае это будет терм в так называемой длинной форме. Далее будет доказан ряд лемм, имеющих отношение к преобразованию $3_{\wedge\eta}$ и проясняющих его смысл.

Лемма 3.1. *Пусть x – переменная, M_i – термы, τ – тип в нормальной форме без пересечений на верхнем уровне (см. следствие 1.5.1). Тогда равносильно:*

$$\begin{aligned} \Gamma \vdash xM_1 \dots M_k : \tau \\ \iff \\ \Gamma \vdash x : \alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau \text{ и } \Gamma \vdash M_i : \alpha_i \text{ для } i = 1 \dots k \end{aligned}$$

Доказательство. Докажем равносильность в обе стороны.

\Leftarrow) применяя правило $(E\rightarrow)$ k раз, получаем в точности необходимое утверждение о типизации.

\Rightarrow) Если $k = 0$, то утверждение леммы очевидно. Иначе докажем индукцией по размеру дерева вывода (анализируя дерево с конца, аналогично доказательству

Леммы 1.7). Рассуждения будет удобнее провести в системе $\lambda_{\wedge \leq} (3)$, поскольку в ней меньше правил вывода.

Какое правило вывода могло быть применено последним, чтобы получить утверждение $\Gamma \vdash xM_1 \dots M_k : \tau$? Это не могло быть (Ax) , так как оно типизирует переменную; это не могло быть $(I \rightarrow)$, так как оно типизирует лямбда абстракцию; это не могло быть $(I \wedge)$, поскольку в τ нет пересечений на верхнем уровне. Таким образом, могли быть применены только два правила: (\leq) или $(E \rightarrow)$.

Согласно Лемме 1.5.1, τ имеет вид $\rho_1 \rightarrow \dots \rightarrow \rho_m \rightarrow \beta$, где β — типовая переменная. По Лемме 1.6, подтип τ тогда обязан иметь вид $\varphi_1 \wedge (\overline{\rho_1} \rightarrow \varphi_2 \wedge (\overline{\rho_2} \rightarrow \varphi_3 \wedge \dots (\overline{\rho_k} \rightarrow \beta) \dots))$. К такой типизации можно или снова применить (\leq) , получив нечто такой же формы (поскольку $\underline{\tau} = \underline{\tau}$) или применить $(I \wedge)$, получив в одной из веток нечто той же формы. Так или иначе, в определённый момент будет применено $(E \rightarrow)$ к $\underline{\tau}$.

Таким образом, в общем случае дерево вывода имеет следующий вид:

$$\frac{\begin{array}{c} \vdots \\ \Gamma \vdash xM_1 \dots M_{k-1} : \alpha_k \rightarrow \underline{\tau} \end{array} \quad \begin{array}{c} \vdots \\ \Gamma \vdash M_k : \alpha_k \end{array}}{\Gamma \vdash xM_1 \dots M_k : \underline{\tau}} (E \rightarrow)$$

$$\begin{array}{c} \vdots \\ \vdots (\leq), (I \wedge) \\ \Gamma \vdash xM_1 \dots M_k : \tau \end{array}$$

Поскольку $\alpha_k \rightarrow \underline{\tau} \leq \alpha_k \rightarrow \tau$, верно утверждение о типизации $\Gamma \vdash xM_1 \dots M_{k-1} : \alpha_k \rightarrow \tau$, которое подходит под условия леммы. Поэтому, применив те же рассуждения ещё $k - 1$ раз, получим $\Gamma \vdash x : \alpha_1 \rightarrow \underline{\alpha_2 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau}$, а также $\Gamma \vdash M_i : \alpha_i$ для $i = 1 \dots k$.

Легко видеть, что $\alpha_1 \rightarrow \underline{\alpha_2 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau} \leq \alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau$, поэтому мы можем применить (\leq) и получить необходимую типизацию.

□

Замечание 3.1.1. Условие о том, что τ — тип без пересечений на верхнем уровне, существенно. Так, например, для $\Gamma = \{x : (\alpha \rightarrow \beta) \wedge (\gamma \rightarrow \delta), M : (\alpha \wedge \gamma)\}$, $\Gamma \vdash xM : \beta \wedge \gamma$, но при этом $\Gamma \not\vdash x : \dots \rightarrow \beta \wedge \gamma$.

Замечание 3.1.2. Условие о том, что τ — тип в нормальной форме, существенно. Так, например, для $\Gamma = \{x : (\alpha \rightarrow \varphi \rightarrow \beta) \wedge (\gamma \rightarrow \varphi \rightarrow \delta), M : (\alpha \wedge \gamma)\}$, $\Gamma \vdash xM : \varphi \rightarrow (\beta \wedge \gamma)$, но при этом $\Gamma \not\vdash x : \dots \rightarrow \varphi \rightarrow (\beta \wedge \gamma)$.

Следствие 3.1.1. В условиях предыдущей леммы, при выполнении любой из равносильных частей, $\Gamma \ni x : \underline{\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau}$,

Доказательство. Для доказательства достаточно провести ещё один шаг в рассуждениях леммы, за тем лишь исключением, что вместо $(E \rightarrow)$ должно быть применено (Ax) , что и гарантирует наличие нужной типизации в контексте. \square

Следствие 3.1.2. *В условиях предыдущей леммы, при выполнении любой из равносильных частей, в контексте Γ есть типизация $x: \sigma$ такая, что $\hat{\sigma}^* \ni \bar{\alpha}_1 \rightarrow \dots \rightarrow \bar{\alpha}_k \rightarrow \tau$.*

В частности, если τ — типовая переменная, то $\hat{\sigma}^ \ni \bar{\alpha}_1 \rightarrow \dots \rightarrow \bar{\alpha}_k \rightarrow \tau$.*

Доказательство. Воспользуемся предыдущим следствием и тем наблюдением, что $\alpha_1 \rightarrow \dots \rightarrow \alpha_k \rightarrow \tau = \varphi_1 \wedge (\bar{\alpha}_1 \rightarrow \varphi_2 \wedge (\bar{\alpha}_2 \rightarrow \varphi_3 \wedge \dots (\bar{\alpha}_k \rightarrow \tau) \dots))$ \square

Следующее следствие проясняет, как с алгоритмической точки зрения устроено преобразование $\mathcal{Z}_{\wedge\eta}$ в Алгоритме 3. А именно, как выбрать x , k и ρ_i^j .

Следствие 3.1.3. *Пусть в шаге $\mathcal{Z}_{\wedge\eta}$ Алгоритма 3, τ_i — типовая переменная. Тогда для того, чтобы выбрать x , k и ρ_i^j ($j = 1 \dots k$), достаточно перебрать все элементы $(x: \sigma)$ контекста такие, что $\hat{\sigma}^*$ содержит тип, заканчивающийся на τ_i , то есть имеющий вид $\rho_i^1 \rightarrow \dots \rightarrow \rho_i^k \rightarrow \tau_i$.*

Для того, чтобы проверить, что выбранный x удовлетворяет всем остальным задачам, а также чтобы подобрать ρ_t^j (для всех $t \neq i$ и $j = 1 \dots k$), достаточно найти среди элементов $\hat{\sigma}^$ (где $\Gamma_t \ni (x: \sigma)$) типы, у которых после «отщепления» k аргументов остаётся подтип τ_t , то есть имеющие вид $\rho_t^1 \rightarrow \dots \rightarrow \rho_t^k \rightarrow \tau_t'$, где $\tau_t' \leq \tau_t$.*

Теорема 2 (Completeness). *Если у системы задач существует решение, то Алгоритм 3 найдёт его или другое решение.*

Доказательство. Доказательство — индукция по размеру наибольшего типа в задачах системы. Пусть дана система $T = [\Gamma_1 \vdash X: \tau_1, \dots, \Gamma_n \vdash X: \tau_n]$, а M — её решение.

Если один из τ_i — тип-пересечение, алгоритм разобьёт его на два. При этом решение M останется решением системы, но уменьшится размер задач системы, поэтому можно применить индукционное предположение. Аналогичные рассуждения можно провести в случае, если один из τ_i не находится в нормальной форме (при этом, чтобы размеры типов уменьшились, нужно снять пересечения, сделав несколько раз преобразование 1).

Пусть ни один из типов τ_i не является пересечением. Рассмотрим, как может быть устроено M .

- $M = \lambda x. M'$.

Тогда все типы в задачах должны быть стрелочные: $\tau_i = \alpha_i \rightarrow \beta_i$. Алгоритм перейдёт к системе $T = [\Gamma_1, x: \alpha_1 \vdash X': \beta_1, \dots, \Gamma_n, x: \alpha_n \vdash X': \beta_n]$, у которой существует решение M' .

- $M = xM_1M_2 \dots M_k$.

- Пусть все типы в задачах стрелочные: $\tau_i = \alpha_i \rightarrow \beta_i$. Тогда алгоритм, аналогично предыдущему пункту, перейдёт к системе $T = [\Gamma_1, y: \alpha_1 \vdash X': \beta_1, \dots, \Gamma_n, y: \alpha_n \vdash X': \beta_n]$. решение которой — $xM_1M_2 \dots M_k y$.
- Пусть один из типов в задачах — переменная: $\tau_i = \beta$. По Лемме 3.1, верно следующее: $\Gamma_i \vdash x: \alpha_i^1 \rightarrow \dots \rightarrow \alpha_i^k \rightarrow \tau_i$ и $\Gamma_i \vdash M_j: \alpha_i^j$ для $i = 1 \dots n, j = 1 \dots k$. Поэтому, выбрав k и x , алгоритм перейдёт к системам $T_j = [(\Gamma_1 \vdash X_j: \overline{\alpha_1^j}), \dots, (\Gamma_n \vdash X_j: \overline{\alpha_n^j})]$. M_j является решением j -й системы.

□

3.3. Завершаемость

В предыдущих разделах было доказано, что Алгоритм 3 выдаёт корректное решение, если оно существует. Однако от алгоритма требуется, чтобы он завершался за конечное число шагов, даже если тип пустой. В текущей системе типов такие гарантии получить невозможно: задача обитаемости неразрешима [11]. Для обеспечения завершаемости можно ограничить множество входных типов, а именно ввести ограничение на их ранг.

Понятие ранга вводится в [8] и определяется как максимальная глубина вложенности « \wedge » в качестве левого аргумента « \rightarrow ». Более формальное определение следующее:

Определение 6.

$$rank(\tau) = 0$$

если τ — тип без пересечений

$$rank(\sigma \wedge \tau) = \max(1, rank(\sigma), rank(\tau))$$

$$rank(\sigma \rightarrow \tau) = \max(1 + rank(\sigma), rank(\tau))$$

если $rank(\sigma) > 0$ или $rank(\tau) > 0$

Так, например, $rank(\varphi \wedge (\tau \rightarrow (\alpha \wedge \beta) \rightarrow \gamma)) = 1$

Лемма 3.2. Если $rank(\rho_1 \rightarrow \dots \rightarrow \rho_k \rightarrow \tau) = r > 0$, то $rank(\rho_i) < r$ для всех $i = 1 \dots k$.

Следующее доказательство во многом повторяет доказательство завершаемости из [7].

Теорема 3 (Termination). Пусть τ — тип с рангом не больше двух ($rank(\tau) \leq 2$). Тогда Алгоритм 3 завершается на входе τ .

Доказательство. Пусть $\text{rank}(\tau) = 0$, что равносильно тому, что в τ нет пересечений. Тогда алгоритм завершится после линейного числа преобразований 2 (и затем одного $3_{\wedge\eta}$).

Проследим за рангами целевых типов τ_i и типов в контекстах Γ_i .

В преобразованиях 0 и 1 контексты не изменяются, а ранги типов в задачах не увеличиваются.

В преобразовании 2 ранг типов в задачах не увеличивается. А в контексты попадают типы σ_i , находящиеся слева от стрелки в типе $\tau_i = \sigma_i \rightarrow \rho_i$, а значит, имеющие или нулевой ранг, или хотя бы на 1 меньший ранг, чем τ_i (см. Лемму 3.2).

В преобразовании $3_{\wedge\eta}$ контексты не изменяются, а в целевыми становятся типы α_i^j , являющиеся аргументами в типах из контекстов, а значит, имеющие на 1 меньший ранг (или нулевой).

Из вышеизложенных соображений следует, что типы в контекстах всегда имеют ранг ≤ 1 , поэтому преобразование $3_{\wedge\eta}$ порождает задачи, в которых целевые типы имеют нулевой ранг. Эти задачи решаются за линейное число шагов. Преобразования 0 – 2 структурно уменьшают целевые типы, поэтому применяются лишь конечное число раз.

□

3.4. Выводы и результаты по главе

В данной главе для Алгоритма 3 были показаны корректность (Теорема 1) и полнота (Теорема 2). Кроме того, была доказана важная Лемма 3.1, использующаяся в доказательстве полноты, и Следствия (3.1.1, 3.1.2 и 3.1.3) из неё. Эти следствия позволяют реализовать преобразование $3_{\wedge\eta}$, явным образом перебирая головную переменную в контексте. Также было введено понятие ранга типа и доказана завершаемость алгоритма для типов, имеющих ранг ≤ 2 .

Таким образом, в главе доказано, что задача обитаемости является разрешимой в системе $\lambda_{\wedge\eta}$ для типов ранга ≤ 2 .

Список литературы

- [1] Alessi Fabio, Barbanera Franco, Dezani-Ciancaglini Mariangiola. Intersection Types and Lambda Models. — Vol. 355, no. 2. — P. 108–126. — Access mode: <https://doi.org/10.1016/j.tcs.2006.01.004>.
- [2] Barendregt Henk, Dekkers Wil, Statman Richard. Lambda Calculus with Types. — New York, NY, USA : Cambridge University Press. — ISBN: 0-521-76614-1 978-0-521-76614-2.
- [3] Coppo Mario, Dezani-Ciancaglini Mariangiola. An Extension of the Basic Functionality Theory for the (λ) -Calculus. — Vol. 21, no. 4. — P. 685–693. — Access mode: <https://doi.org/10.1305/ndjfl/1093883253>.
- [4] Hindley J. R. The Simple Semantics for Coppo-Dezani-Sallé Types // International Symposium on Programming / Ed. by Mariangiola Dezani-Ciancaglini, Ugo Montanari. — Berlin, Heidelberg : Springer Berlin Heidelberg. — P. 212–226.
- [5] Hindley J. Roger. Types with Intersection: An Introduction. — Vol. 4, no. 5. — P. 470–486. — Access mode: <https://doi.org/10.1007/BF01211394>.
- [6] Kfoury A. J., Wells J. B. Principality and Type Inference for Intersection Types Using Expansion Variables. — Vol. 311, no. 1-3. — P. 1–70. — Access mode: <https://doi.org/10.1016/j.tcs.2003.10.032>.
- [7] Kuśmierek Dariusz. The Inhabitation Problem for Rank Two Intersection Types // Typed Lambda Calculi and Applications / Ed. by Simona Ronchi Della Rocca. — Berlin, Heidelberg : Springer Berlin Heidelberg. — P. 240–254.
- [8] Leivant Daniel. Polymorphic Type Inference // Proceedings of the 10th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages. — POPL '83. — New York, NY, USA : ACM. — P. 88–98. — Access mode: <http://doi.acm.org/10.1145/567067.567077>.
- [9] Neergaard Peter Møller, Mairson Harry G. Types, Potency, and Idempotency: Why Nonlinearity and Amnesia Make a Type System Work // Proceedings of the Ninth ACM SIGPLAN International Conference on Functional Programming, ICFP 2004, Snow Bird, UT, USA, September 19-21, 2004. — P. 138–149. — Access mode: <https://doi.org/10.1145/1016850.1016871>.
- [10] Pottinger Garrel. A Type Assignment for the Strongly Normalizable (λ) -Terms // J.R. Hindley; J.P. Seldin (Eds.), To H. B. Curry : Essays on Combinatory Logic, Lambda Calculus and Formalism. — London : Academic Press. — P. 561–577.

- [11] Urzyczyn Pawel. The Emptiness Problem for Intersection Types. — Vol. 64, no. 3. — P. 1195–1215. — jstor : 2586625.
- [12] Urzyczyn Pawel. Inhabitation in Typed Lambda-Calculi (A Syntactic Approach) // Typed Lambda Calculi and Applications, Third International Conference on Typed Lambda Calculi and Applications, TLCA '97, Nancy, France, April 2-4, 1997, Proceedings / Ed. by Philippe de Groote. — Vol. 1210 of Lecture Notes in Computer Science. — Springer. — P. 373–389.
- [13] Urzyczyn Paweł. Inhabitation of Low-Rank Intersection Types // Proceedings of the 9th International Conference on Typed Lambda Calculi and Applications. — TLCA '09. — Berlin, Heidelberg : Springer-Verlag. — P. 356–370. — Access mode: http://dx.doi.org/10.1007/978-3-642-02273-9_26.
- [14] Wells Joe B., Haack Christian. Branching Types // Programming Languages and Systems, 11th European Symposium on Programming, ESOP 2002, Held as Part of the Joint European Conference on Theory and Practice of Software, ETAPS 2002, Grenoble, France, April 8-12, 2002, Proceedings. — P. 115–132. — Access mode: https://doi.org/10.1007/3-540-45927-8_9.