

UNIVERSIDADE DO ESTADO DE SANTA CATARINA — UDESC
CENTRO DE CIÊNCIAS TECNOLÓGICAS — CCT
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO — BCC

VINÍCIOS BIDIN SANTOS

INFERÊNCIA DE TIPOS PARA CPS

JOINVILLE

2025

VINÍCIOS BIDIN SANTOS

INFERÊNCIA DE TIPOS PARA CPS

Trabalho de conclusão de curso submetido à Universidade do Estado de Santa Catarina como parte dos requisitos para a obtenção do grau de Bacharel em Ciência da Computação

Orientador: Cristiano Damiani Vasconcellos

Coorientador: Paulo Henrique Torrens

JOINVILLE

2025

Para gerar a ficha catalográfica de teses e
dissertações acessar o link:
<https://www.udesc.br/bu/manuais/ficha>

Santos, Vinícios Bidin
Inferência de tipos para CPS / Vinícios Bidin Santos.
-- Joinville, 2025.
65 p. il.; 30 cm.

Orientador: Cristiano Damiani Vasconcellos.
Coorientador: Paulo Henrique Torrens.
Trabalho de Conclusão de Curso - Universidade
do Estado de Santa Catarina, Centro de Ciências
Tecnológicas, Bacharelado em Ciência da Computação,
Joinville, 2025.

1. Inferência de Tipos. 2. Estilo de Passagem de
Continuação. 3. Algoritmo W. 4. Damas-Milner. 5. Haskell.
I. Vasconcellos, Cristiano Damiani . II. Torrens, Paulo
Henrique . III. Universidade do Estado de Santa Catarina,
Centro de Ciências Tecnológicas, Bacharelado em Ciência
da Computação. IV. Título.

VINÍCIOS BIDIN SANTOS

INFERÊNCIA DE TIPOS PARA CPS

Trabalho de conclusão de curso submetido à Universidade do Estado de Santa Catarina como parte dos requisitos para a obtenção do grau de Bacharel em Ciência da Computação

Orientador: Cristiano Damiani Vasconcellos

Coorientador: Paulo Henrique Torrens

BANCA EXAMINADORA:

Orientador:

Dr. Cristiano Damiani Vasconcellos
UDESC

Coorientador:

Me. Paulo Henrique Torrens
University of Kent

Membros:

Dra. Karina Girardi Roggia
UDESC

Me. Gabriela Moreira
UDESC

Joinville, Julho de 2025

A todos que me apoiaram nesses anos!

AGRADECIMENTOS

Quero deixar registrados aqui os meus mais profundos agradecimentos a todos que estiveram ao meu lado durante este período. Tanto nos momentos de alegria, compartilhando conquistas e bons momentos, quanto nas fases difíceis e de frustração, oferecendo apoio e ouvindo quando mais precisei.

Um abraço apertado à minha família – mãe, pai, irmãos – e à minha namorada, que estiveram sempre por perto, torcendo por mim e dividindo cada passo dessa caminhada.

Agradeço também aos ótimos amigos que fiz nesse caminho, entre colegas e professores. Em especial, ao pessoal dos grupos do Função e do Suporte, onde não faltaram boas risadas e valiosos aprendizados.

No final, é sobre as pessoas com quem decidimos dividir os momentos preciosos. E nesse ponto, sou muito sortudo pois estou cercado delas!

“Eu sou a senhora Marocas.”

(Senhora Marocas — Toy Story, [1995])

RESUMO

No contexto de compiladores, as representações intermediárias desempenham um papel fundamental, especialmente em ambientes de produção. O *Continuation Passing Style* (CPS) é uma dessas representações, notável para linguagens funcionais devido às otimizações avançadas que permite. Contudo, há uma escassez de pesquisas e implementações dessa IR com termos devidamente tipados, o que impede a detecção de uma gama de erros durante as etapas de compilação. Este trabalho, portanto, explora a formalização de um sistema de tipos para o CPS e o desenvolvimento de um algoritmo de inferência de tipos para essa representação, juntamente com sua implementação na linguagem de programação Haskell.

Palavras-chave: Inferência de Tipos, Estilo de Passagem de Continuação (CPS), Damas-Milner, Haskell, Sistema de Tipos.

ABSTRACT

In the field of compilers, intermediate representations (IR) play a fundamental role, particularly in production environments. Continuation Passing Style (CPS) is one such representation, notable for functional languages due to the advanced optimizations it enables. However, research and implementations of CPS with well-typed terms remains limited, which restricts the detection of certain classes of errors during compilation stages. This work investigates the formalization of a type system for CPS and develops a type inference algorithm for this representation, along with its implementation in the Haskell programming language.

Keywords: Type Inference, Continuation Passing Style (CPS), Damas-Milner, Haskell, Type System.

LISTA DE ILUSTRAÇÕES

Figura 2.1 – Sequência de representações intermediárias	17
Figura 2.2 – Regras de aplicação da substituição de tipos	27
Figura 2.3 – Regras de Inferência do sistema Damas-Milner	28
Figura 2.4 – Algoritmo de unificação para o Sistema Damas-Milner no formato de função.	29
Figura 2.5 – Algoritmo de verificação de ocorrência para o Sistema Damas-Milner no formato de função.	30
Figura 2.6 – Algoritmo W em formato de regras de inferência.	30
Figura 3.1 – Algoritmo de unificação para o Sistema baseado em continuações no formato de função.	38

LISTA DE CÓDIGOS

2.1	Função fatorial em Haskell	18
2.2	Função fatorial em Haskell com chamada de cauda	18
2.3	Função soma em Haskell em Estilo Direto	20
2.4	Função soma em Haskell em CPS	20
2.5	Função fatorial em Haskell em Estilo Direto	21
2.6	Função fatorial em Haskell em CPS	21
2.7	Função somatório de elementos de lista em Haskell	24
3.1	Definição dos tipos de dados	39
3.2	Continuação inicial	40
3.3	Tradução das expressões para CBN	40
3.4	Tradução da função identidade em CBN	41
3.5	Tradução do numeral de Church “2” em CBN	42
3.6	Tradução das expressões para CBV	42
3.7	Tradução da função identidade em CBV	43
3.8	Tradução do numeral de Church “2” em CBV	43
3.9	Tradução dos tipos para CBN	44
3.10	Tradução dos tipos para CBV	45
3.11	Função principal de Inferência	45
3.12	Verificação de Subtipagem	47
3.13	Geração de código para computação de numerais de Church	49
3.14	Tradução do numeral de Church “0” em CBN	50
3.15	Tradução do numeral de Church “0” em CBV	50
3.16	Código gerado ao traduzir o numeral de Church “0”	51
3.17	Execução do programa principal	52
3.18	Execução do programa gerado	53
4.1	Tradução em CBN da identidade com ‘let’	56
4.2	Tradução em CBV da identidade com ‘let’	56
4.3	Tradução em CBN da autoaplicação	57
4.4	Tradução em CBV da autoaplicação	58
4.5	Tradução em CBN do combinador Y	58
4.6	Tradução em CBV do combinador Y	60

LISTA DE ABREVIATURAS E SIGLAS

IR	Intermediate Representation
CPS	Continuation Passing Style
SSA	Static Single Attribution
ANF	A Normal Form
CBN	Call by Name
CBV	Call by Value

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVO GERAL	14
1.2	OBJETIVOS ESPECÍFICOS	14
1.3	TRABALHOS RELACIONADOS	14
1.4	METODOLOGIA	15
1.5	ESTRUTURA DO TRABALHO	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	REPRESENTAÇÃO INTERMEDIÁRIA DE CÓDIGO	17
2.1.1	CPS	17
2.2	TEORIA DE TIPOS	22
2.3	CÁLCULO LAMBDA SIMPLEMENTE TIPADO	24
2.4	SISTEMA DAMAS-MILNER	25
2.4.1	Algoritmo W	29
2.5	CÁLCULO DE CPS SIMPLEMENTE TIPADO	31
2.5.1	Tradução CPS	33
3	DESENVOLVIMENTO	35
3.1	FORMALIZAÇÃO	35
3.2	IMPLEMENTAÇÃO	39
3.2.1	Tipos de Dados	39
3.2.2	Traduções	40
3.2.3	Inferência	45
3.2.4	Geração de Código	49
3.2.5	Fluxo Principal	52
4	RESULTADOS	54
4.1	COMBINADOR S	54
4.2	SOMA	54
4.3	MULTIPLICAÇÃO	55
4.4	IDENTIDADE COM ‘LET’	56
4.5	AUTOAPLICAÇÃO MAL-TIPADA	57
4.6	COMBINADOR Y	58
5	CONCLUSÃO	62
5.1	CONSIDERAÇÕES FINAIS	62
	REFERÊNCIAS	64

1 INTRODUÇÃO

A compilação de programas envolve diversas fases, cada uma com funções específicas, como análise léxica, análise sintática, análise semântica, otimizações, e, finalmente, a geração de código. Uma etapa crítica nesse processo é a otimização, que frequentemente se baseia em representações intermediárias (IRs). Essas representações atuam como ponte entre o código fonte e o código de máquina, permitindo que transformações e otimizações sejam aplicadas de maneira mais eficaz (PLOTKIN, 1975).

As representações intermediárias variam conforme o paradigma da linguagem de programação. Para linguagens imperativas, a Representação em Atribuição Única Estática (SSA) é amplamente adotada. Já em linguagens funcionais, a Forma Normal A (ANF) e o Estilo de Passagem de Continuação (CPS) se destacam. Este trabalho foca especificamente no CPS, uma IR que oferece vantagens particulares em termos de otimização e simplicidade na geração de código.

Essas características do CPS se tornam ainda mais evidentes quando é comparado como diferentes linguagens lidam com o fluxo de execução. Em linguagens de alto nível, por exemplo, a pilha de chamadas atua como uma abstração fundamental para gerenciar o controle de retorno das funções. No entanto, em linguagens de baixo nível, como *assembly*, o controle de fluxo é mais explícito e depende do uso direto de instruções, que operam sobre a pilha de chamadas para empilhar e desempilhar os endereços de retorno. Nesse contexto, o CPS se destaca ao tornar as continuações explicitamente representadas no código. Em vez de confiar na pilha de chamadas para gerenciar retornos, o CPS introduz um parâmetro adicional em cada função, representando a continuação — isto é, o que deve ser feito com o resultado da função (KENNEDY, 2007). Desta forma, em vez de simplesmente retornar um valor diretamente, a função invoca essa continuação, transferindo explicitamente o controle à próxima etapa da computação. Isso elimina a dependência da pilha de chamadas, simplificando o modelo de execução e tornando-o mais alinhado com as necessidades de linguagens de baixo nível.

Além disso, a adoção do CPS como representação intermediária vai além da tradução de linguagens de alto nível para código de máquina. O CPS facilita a aplicação de otimizações avançadas, como a eliminação de chamadas de cauda e a fusão de funções, além de permitir uma correspondência mais direta com o código gerado em linguagens de montagem (FLANAGAN et al., 1993).

Por outro lado, um ponto importante a ser considerado é que, apesar de existirem sistemas de tipos para CPS — como o sistema simplesmente tipado proposto por Thielecke (1997) — muitas implementações optam por representações não tipadas (MORRISETT et al., 1999; TORRENS; ORCHARD; VASCONCELLOS, 2024). Apesar da predominância de linguagens intermediárias não tipadas, pode ser citado como exceção o projeto LLVM — uma infraestrutura amplamente utilizada para construção de compiladores — que adota uma IR com suporte a um sistema de tipos próprio. Esse suporte é ainda expandido com o uso do MLIR, uma extensão

modular da LLVM voltada à construção de representações intermediárias com estruturas de tipos mais expressivas e flexíveis. Ainda assim, na prática, é comum que as IRs não estendam completamente os tipos até essa etapa da compilação.

Embora essa abordagem simplifique a implementação inicial, ela pode comprometer a segurança e a correção do código. Um sistema de tipos robusto pode não apenas garantir a correção de certas transformações e otimizações, mas também identificar uma classe inteira de erros antes da execução, proporcionando assim maior confiabilidade ao processo de compilação.

Diante dessas considerações, este trabalho propõe apresentar e desenvolver uma formalização de um sistema de tipos para CPS, bem como um algoritmo de inferência de tipos para o mesmo. A escolha da linguagem de programação para a solução proposta será Haskell. Por ser uma linguagem funcional pura fortemente tipada, possui características desejáveis como transparência referencial (SØNDERGAARD; SESTOFT, 1990) e um sistema de tipos robusto para explorar as vantagens do CPS e aplicar o sistema de tipos de maneira rigorosa. Dessa forma, a escolha de Haskell não apenas facilita o desenvolvimento de uma implementação segura e eficiente do CPS, como também conta com garantias de seguranças que são fundamentais para o sucesso deste trabalho.

1.1 OBJETIVO GERAL

Este trabalho tem como objetivo formalizar uma extensão para o sistema de tipos para CPS proposto por Thielecke (1997), adicionando polimorfismo com suporte a inferência para esta representação intermediária.

1.2 OBJETIVOS ESPECÍFICOS

- Formalizar um sistema de tipos para CPS com suporte a polimorfismo;
- Propor e implementar em Haskell um algoritmo de inferência de tipos para CPS;
- Validar a implementação do algoritmo por meio do teste de inferência para expressões. Se possível, realizar a geração de programas para verificação de que o algoritmo infere corretamente os tipos a eles.

1.3 TRABALHOS RELACIONADOS

Após um estudo na literatura a respeito da tipagem de linguagens de mais baixo nível em relação ao código fonte, alguns trabalhos correlatos foram encontrados.

Em (NECULA, 1997), o autor introduz o conceito de um código que carrega provas (PCC do inglês *Proof-Carrying Code*), um paradigma em que programas de código de máquina carregam junto a si provas formais de que obedecem a certas propriedades de segurança. O sistema de verificação do consumidor do código pode então validar essas provas de forma

automática, sem necessidade de confiar no produtor. O modelo proposto utiliza uma linguagem de políticas de segurança, uma lógica para especificações e provas, e uma máquina abstrata como base formal. O trabalho demonstra a viabilidade prática do modelo com a implementação de um verificador rápido para código de baixo nível, destacando a aplicabilidade do PCC em ambientes com execução de código não confiável, como acontece em sistemas distribuídos.

Em (MORRISETT et al., 1999), os autores propõem uma linguagem de montagem tipada (TAL do inglês *Typed Assembly Language*) e um compilador que traduz programas do Sistema F (do inglês *System F*) para TAL de forma preservadora de tipos. A linguagem TAL é baseada em uma arquitetura reduzida (RISC do inglês *Reduced Instruction Set Computer*) e possui um sistema de tipos capaz de representar abstrações de linguagens de alto nível, como fechamentos e tuplas. A tradução é dividida em etapas como conversão para CPS e conversão de fechamentos, mantendo a tipagem correta em cada fase. O trabalho também apresenta uma abordagem simplificada para conversão de fechamentos polimórficos e discute o uso de TAL para geração de código seguro em ambientes com código não confiável.

Em (SHAO; LEAGUE; MONNIER, 1998), os autores descrevem a implementação da linguagem intermediária tipada FLINT no compilador SML/NJ, com foco na eficiência e escalabilidade de compiladores que preservam tipos. O trabalho defende que um compilador que mantém informações de tipo não será escalável a menos que todas as suas fases também preservem a complexidade assintótica de tempo e espaço ao representar e manipular tipos. Para isso, são combinadas técnicas como *hash-consing*, memoização e codificação lambda com suspensões, garantindo representações compactas em forma de grafos acíclicos direcionados (DAGs do inglês *Directed Acyclic Graphs*). Os autores mostram que, com essas técnicas, é possível realizar operações como substituição, igualdade e redução de tipos de forma eficiente, o que viabiliza o uso prático de linguagens intermediárias fortemente tipadas.

Em (BOWMAN et al., 2018), os autores demonstram que a tradução CPS preservadora de tipos para linguagens dependentemente tipadas com tipos Σ e Π não é impossível, como sugerido por resultados anteriores. Utilizando *answer-type polymorphism*, eles propõem novas traduções CPS *call-by-name* e *call-by-value* a partir do Cálculo de Construções, provando preservação de tipos e correção da compilação. Para isso, estendem a linguagem alvo com regras adicionais justificadas por um teorema livre, e provam a consistência desse sistema por meio de um modelo no Cálculo das Construções extensional, obtendo assim segurança em tempo de ligação (do inglês *link-time safety*).

1.4 METODOLOGIA

A metodologia deste trabalho consistiu em duas principais etapas: pesquisa bibliográfica e implementação. A primeira envolveu uma extensa revisão de literatura sobre continuações e seu cálculo, bem como um aprofundamento no estudo de sistemas de tipos, com o objetivo de proporcionar uma compreensão completa. A segunda contemplou a formalização do sistema

de tipos e do algoritmo de inferência para o cálculo de continuações, junto da implementação destes.

No escopo deste trabalho, a validação do algoritmo ocorreu por meio de testes de implementação, analisando os tipos inferidos das expressões. Em etapa posterior, serão necessárias as provas de consistência e de completude do algoritmo em relação ao sistema de tipos proposto.

1.5 ESTRUTURA DO TRABALHO

A primeira etapa do trabalho consistiu principalmente na fundamentação teórica e revisão bibliográfica no estudo de CPS e sistemas de tipos. Em razão disto, o Capítulo 2 contém os conceitos e definições necessários para entendimento do tema. Este é separado em seções, tal que a Seção 2.1 aborda representação intermediária, com um aprofundamento em CPS na Subseção 2.1.1. Teoria de tipos é então apresentada na Seção 2.2, detalhando o Cálculo Lambda Simplesmente Tipado na Seção 2.3 Um aprofundamento no sistema Damas-Hindley-Milner na Seção 2.4, discutindo de maneira mais específica o algoritmo W na Subseção 2.4.1.

Já a segunda etapa do trabalho, consistiu da parte prática do trabalho, envolvendo a formalização do sistema de tipos bem como seu desenvolvimento e por fim a conclusão do trabalho. O Capítulo 3 agrupa tanto a formalização do sistema proposto, presente na Seção 3.1 quanto a implementação deste, apresentada em detalhes na Seção 3.2. Os resultados obtidos são então apresentados e discutidos no Capítulo 4, dividindo-o em Seções referentes aos testes realizados. Por fim, o Capítulo 5 conta com considerações finais bem como as conclusões do autor.

É importante destacar que todas as formalizações do sistema de tipos apresentadas na Seção 3.1 foram elaboradas pelos orientadores em conjunto com o autor, durante reuniões de desenvolvimento conceitual, explicando as motivações e os raciocínios para chegar no resultado. Até a data de conclusão deste trabalho, tais formalizações ainda não foram publicadas, estando em preparação uma produção acadêmica que as reunirá para futura referência.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 REPRESENTAÇÃO INTERMEDIÁRIA DE CÓDIGO

Um compilador é um programa responsável por traduzir um código escrito em uma linguagem de programação para outra, geralmente do código-fonte para código de máquina, permitindo assim a execução do programa. Durante esse processo, é fundamental que o mínimo de informações seja perdido, uma vez que a semântica original deve ser preservada no processo de tradução. Uma abordagem comum utilizada para manter a integridade semântica e possibilitar otimizações, são as representações intermediárias (IR, do inglês *intermediate representation*) (COOPER; TORCZON, 2014).

Compiladores modernos, amplamente utilizados na indústria, empregam mais de uma IR para tirar proveito das vantagens de cada uma, uma vez que essas representações são projetadas para diferentes objetivos, como otimizações específicas. As IRs podem ser classificadas de acordo com o nível de abstração e são comumente aplicadas em sequência. Representações com um nível maior de abstração são usadas próximas ao código-fonte, enquanto aquelas de nível mais baixo estão mais próximas do código de máquina (AHO et al., 2008), como ilustrado no Figura 2.1.

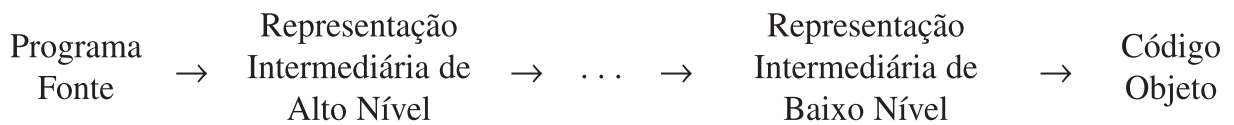


Figura 2.1 – Sequência de representações intermediárias

Fonte: (AHO et al., 2008)

Uma das principais informações que deve ser preservada em uma IR é o fluxo de controle, isto é, a ordem em que as instruções do programa são executadas, como chamadas de função, *loops* e condições. Para garantir que o compilador mantenha a semântica correta do programa, o fluxo de controle deve ser repassado de alguma maneira durante o processo de tradução (COOPER; TORCZON, 2014). Uma das maneiras disso ser feito explicitamente é com o uso de continuações, que são funções que descrevem o próximo passo de uma computação em um ponto particular da execução do programa.

2.1.1 CPS

O estilo de passagem de continuações (CPS, do inglês *continuation passing style*) é uma técnica de transformação de código que torna o fluxo de controle de um programa explícito, ao converter o estilo convencional de chamadas de função em chamadas que passam explicitamente o controle para a próxima etapa, conhecida como continuação (do inglês, *continuation*) (APPEL, 1992). Em vez de retornar diretamente o resultado de uma função, o CPS transforma cada função para que, ao finalizar sua computação, ela invoque uma continuação que representa o próximo

passo a ser executado no programa. Assim, toda chamada de função se torna uma chamada de cauda.

Uma chamada de cauda (do inglês *tail call*) ocorre quando a última instrução executada em uma função é uma chamada a outra função, sem que restem computações adicionais a serem feitas após essa chamada (MUCHNICK, 1997). Isso permite que a função atual libere seu quadro de ativação, otimizando o uso de memória, já que o compilador não precisa manter o estado da função anterior na pilha. Em contraste, uma chamada que não é de cauda ocorre quando ainda restam operações após a chamada, como somas ou multiplicações, o que exige que o quadro de ativação da função atual permaneça na pilha até a conclusão dessas operações.

No Código 2.1, o exemplo da função fatorial demonstra uma chamada que não é de cauda, pois a chamada recursiva `factorial (n - 1)` não é a última operação a ser realizada. A função precisa aguardar o retorno desta chamada para, então, multiplicar o resultado por `n`, o que impede a liberação do quadro de ativação até o término da multiplicação.

Em contraste, no Código 2.2 como exemplo, tem-se uma versão da função fatorial que utiliza chamada de cauda. A função auxiliar `go` acumula o valor do cálculo diretamente em seu argumento `a`, e a chamada recursiva `go (n - 1) (a * n)` é a última instrução a ser executada. Como não há operações pendentes após a chamada recursiva, o compilador pode otimizar a função, reutilizando o espaço reservado para o quadro de ativação da função `go` para a chamada subsequente, tornando o cálculo mais eficiente.

```
1 factorial :: Int -> Int
2 factorial 0 = 1
3 factorial n = n * factorial (n - 1)
```

Código 2.1 – Função fatorial em Haskell

```
1 go :: Int -> Int -> Int
2 go 1 a = a
3 go n a = go (n - 1) (a * n)
4
5 factorial :: Int -> Int
6 factorial 0 = 1
7 factorial n = go n 1
```

Código 2.2 – Função fatorial em Haskell com chamada de cauda

O cálculo lambda, definido por Church (1932), é um sistema formal que serve como base para a maioria das linguagens funcionais. Ele é capaz de representar qualquer computação utilizando abstrações e aplicações através de reduções. Sua sintaxe consiste em três regras simples que definem os elementos principais do sistema: variável, abstração e aplicação, conforme apresentados a seguir:

$$\text{Termo } e ::= x \mid \lambda x. e \mid e e$$

A partir dessa sintaxe, um termo e pode possuir apenas uma das três formas. A primeira forma refere-se às variáveis, que representam identificadores no sistema. A segunda forma, chamada de abstração, define uma função lambda: uma função que associa o identificador x a um termo e , seu corpo, com x vinculado ao termo e . Finalmente, a aplicação ocorre quando um termo e é aplicado a outro e , representando a chamada de uma função.

No cálculo lambda, as variáveis podem ser classificadas como livres ou ligadas, dependendo de seu contexto em um termo. Variáveis são consideradas livres quando não estão associadas a uma abstração de função. Por exemplo, no termo $\lambda x.y$, a variável y é livre, pois não está ligada a nenhum parâmetro introduzido. Em contraste, no termo $(\lambda x.x)y$, a variável x está ligada dentro do corpo da abstração, enquanto y permanece livre. Um termo sem variáveis livres é denominado fechado ou combinador; por exemplo, $\lambda x.\lambda y.xy$ é um combinador, pois todas as variáveis estão ligadas às suas respectivas abstrações.

Para avaliar expressões no cálculo lambda, usamos três tipos de redução: α , β e η , que seguem as seguintes definições:

α -redução: Renomeação de variáveis ligadas.

$$\lambda x.e \rightarrow \lambda y.e[y/x]$$

Note que, $e[y/x]$ indica a substituição de todas as ocorrências ligadas de x por y em e , desde que y não ocorra livre em e . Por exemplo, $\lambda x.x + z \rightarrow_{\alpha} \lambda y.y + z$.

β -redução: Aplicação de função.

$$(\lambda x.e_1)e_2 \rightarrow e_1[e_2/x]$$

Aqui, $e_1[e_2/x]$ denota a substituição de todas as ocorrências ligadas de x em e_1 por e_2 . Para evitar conflitos com variáveis livres em e_2 , aplica-se α -redução prévia. Por exemplo, em $(\lambda x.\lambda y.x + y)(y + 1)$, renomeia-se y para z na abstração interna:

$$(\lambda x.\lambda z.x + z)(y + 1) \rightarrow_{\beta} \lambda z.(y + 1) + z.$$

η -redução: Expansão de função.

$$\lambda x.(ex) \rightarrow_{\eta} e \quad \text{se } x \notin \text{FV}(e)$$

Tal que $\text{FV}(e)$ denota as variáveis livres em e . Por exemplo, $\lambda x.(\lambda y.y + 1)x \rightarrow_{\eta} \lambda y.y + 1$.

As reduções são responsáveis pela semântica operacional do cálculo lambda. A α -redução permite a renomeação de variáveis ligadas, enquanto a β -redução descreve a aplicação de funções, substituindo o parâmetro da função por um valor passado como argumento. Por fim, a η -redução lida com a simplificação de funções quando elas aplicam diretamente seu argumento.

A transformação para CPS se baseia nessa estrutura formal. No cálculo lambda tradicional, o fluxo de execução é implícito: as funções são aplicadas e seus resultados são retornados

automaticamente. No entanto, no CPS, o fluxo de controle é explicitamente representado como uma série de chamadas a funções. Cada função, em vez de retornar diretamente um valor, recebe um argumento extra, a continuação, que indica o próximo passo da computação.

Por exemplo, a expressão $\lambda x.x + 1$ no cálculo lambda tradicional retornaria o valor $x + 1$. Ao transformar essa expressão para CPS, ela se torna $\lambda x.\lambda k.k(x + 1)$.

Aqui, k é a continuação que processa o resultado $x + 1$. Ao converter funções para CPS, o fluxo de controle do programa passa a ser gerenciado de forma explícita através de chamadas de função que encadeiam as etapas subsequentes. Essa transformação traz diversas vantagens no contexto de compiladores, pois torna explícitas as informações normalmente implícitas na pilha de execução e, assim, abre espaço para a aplicação de uma série de otimizações (APPEL, 1992).

Em primeiro lugar, o CPS expõe todas as chamadas de cauda, permitindo a eliminação de recursividade em cauda (TCO, do inglês *tail-call optimization*). Essa otimização possibilita que chamadas recursivas sejam realizadas sem o acúmulo de quadros de ativação na pilha, reduzindo o uso de memória e evitando estouros de pilha em programas fortemente recursivos.

Além disso, a natureza explícita das chamadas em CPS favorece a expansão *inline* de funções. Como cada chamada é clara e separada, o compilador pode substituir uma chamada de função por sua definição diretamente, reduzindo o *overhead* de chamadas e potencialmente expondo outras oportunidades de otimização.

A transformação para CPS também facilita a alocação de registradores. Com o fluxo de controle explicitado, o compilador pode prever melhor o tempo de vida das variáveis e organizar os dados de forma que o uso de registradores seja maximizado e o número de acessos à memória minimizado (APPEL, 1992).

Por fim, como todo o fluxo de execução é expresso por chamadas encadeadas, o código gerado permite análises estáticas mais precisas. O compilador pode, por exemplo, detectar com mais facilidade padrões de execução, dependências entre expressões e oportunidades de reordenação ou eliminação de código redundante.

Dessa forma, a conversão para CPS não apenas preserva a semântica da computação original, mas também fornece ao compilador uma estrutura rica e detalhada para aplicar otimizações de maneira eficaz. A tradução para CPS então, converte um código escrito em estilo direto (onde o controle de fluxo é implícito) para o estilo de passagem de continuações (FLANAGAN et al., 1993). A principal ideia por trás dessa transformação é modificar as funções para que elas não retornem um valor diretamente, mas, em vez disso, passem o resultado para uma continuação.

```
1 add :: Int -> Int -> Int
2 add x y = x + y
```

Código 2.3 – Função soma em Haskell em Estilo Direto

```
1 addCps :: Int -> Int -> (Int -> r) -> r
2 addCps x y k = k (x + y)
```

Código 2.4 – Função soma em Haskell em CPS

Por exemplo, o Código 2.3 apresenta um programa na linguagem Haskell que soma dois números no estilo direto, retornando o valor após realizar o cálculo. Já o Código 2.4 mostra um programa equivalente em CPS. Nesta versão, o controle de fluxo do programa é explícito, pois a função k é chamada para processar o resultado da soma dos argumentos.

```

1 sub :: Int -> Int -> Int
2 sub x y = x - y
3
4 mult :: Int -> Int -> Int
5 mult x y = x * y
6
7 factorial :: Int -> Int
8 factorial 0 = 1
9 factorial n = mult n (factorial (sub n 1))

```

Código 2.5 – Função fatorial em Haskell em Estilo Direto

```

1 subCps :: Int -> Int -> (Int -> r) -> r
2 subCps x y k = k (x - y)
3
4 multCps :: Int -> Int -> (Int -> r) -> r
5 multCps x y k = k (x * y)
6
7 factorialCps :: Int -> (Int -> r) -> r
8 factorialCps 0 k = k (1)
9 factorialCps n k =
10   subCps n 1 (\nMinus1 ->
11     factorialCps nMinus1 (\factNMinus1 ->
12       multCps n factNMinus1 k))

```

Código 2.6 – Função fatorial em Haskell em CPS

Para ilustrar melhor, o Código 2.5 apresenta um programa em Haskell que calcula o fatorial no estilo direto, utilizando funções definidas para multiplicação e subtração. No Código 2.6, um programa similar em CPS é definido, com as funções auxiliares também transformadas para CPS. Na função `factorialCps` é possível notar duas funções lambda (continuações anônimas): a primeira com parâmetro `nMinus1` e a segunda com parâmetro `factNMinus1`. A primeira continuação guarda o resultado da operação $n - 1$, enquanto a segunda recebe recursivamente o cálculo do fatorial de $n - 1$, multiplica por n e finalmente passa o resultado para a continuação k . Note que aqui, quando `subCps` retorna, seu resultado é vinculado a `nMinus1`; quando `factorialCps` retorna, seu resultado é vinculado a `factNMinus1`. As variáveis n e k são capturadas do contexto externo pelas *closures*, mantendo seu valor original.

Outro fato importante a ser observado nos códigos apresentados é a tipagem das funções. Na função de soma, definida no Código 2.3, a função tem tipo $Int \rightarrow Int \rightarrow Int$, ou seja, ela recebe dois inteiros e retorna um inteiro. Já a função de soma em CPS, definida no Código 2.4, possui o tipo $Int \rightarrow Int \rightarrow (Int \rightarrow r) \rightarrow r$. Isso significa que a função recebe dois inteiros e uma continuação, que é uma função de tipo $Int \rightarrow r$, onde r pode ser qualquer tipo, e retorna esse mesmo tipo r .

Essa transformação de tipo reflete a diferença fundamental entre o estilo direto e o CPS: em vez de retornar um valor diretamente, a função em CPS recebe uma continuação que especifica o próximo passo da computação. O mesmo padrão pode ser observado nas funções para o cálculo do fatorial nos Códigos 2.5 e 2.6. No estilo direto, a função `factorial` tem o tipo $Int \rightarrow Int$, enquanto na versão CPS, a função `factorialCps` tem o tipo $Int \rightarrow (Int \rightarrow r) \rightarrow r$.

Essa correspondência entre os tipos não é uma coincidência. Como discutido por Torrens (2019), uma função em estilo direto com tipo $A \rightarrow B$ pode ser transformada em uma função em CPS com o tipo $A \rightarrow (B \rightarrow \perp) \rightarrow \perp$. Aqui, \perp representa o tipo dos valores que nunca retornam, uma característica associada ao estilo de passagem de continuações, onde as funções são compostas de forma a encadear continuações até que a execução termine de maneira explícita.

Este exemplo simples da função fatorial em CPS ilustra as dificuldades inerentes ao uso de continuações explícitas, como a verbosidade do código e a complexidade de compreensão, mas não sua inadequação para compilação. Apesar desses desafios, o CPS se mostra extremamente adequado para a aplicação de otimizações, sendo uma escolha eficiente para representações intermediárias, especialmente em cenários onde o desempenho é essencial.

2.2 TEORIA DE TIPOS

A Teoria de Tipos, conforme apresentada por (COQUAND, 2022), foi introduzida por Russell em 1908 ao encontrar um paradoxo na Teoria de Conjuntos, conhecido atualmente como o Paradoxo de Russell:

$$\text{Seja } R = \{x \mid x \notin x\}, \text{ então } R \in R \iff R \notin R$$

Ou seja, considere R como o conjunto dos conjuntos que não contêm a si mesmos. A contradição surge ao observar que, se o conjunto R contém a si mesmo, isso implica que R não contém a si mesmo, e vice-versa.

Outra maneira de descrever esse paradoxo é através do Paradoxo do Barbeiro: imagine uma cidade com apenas um barbeiro, onde ele somente barbeia aqueles que não se barbeiam. O paradoxo surge quando perguntamos: “Quem barbeia o barbeiro?” Ele não pode fazer sua própria barba, pois barbeia apenas aqueles que não fazem a própria barba. No entanto, se ele não faz sua própria barba, então pertence ao grupo daqueles que devem ser barbeados pelo barbeiro, logo, ele deveria barbear-se. Essa situação gera uma contradição semelhante ao Paradoxo de Russell.

Assim como os paradoxos na Teoria de Conjuntos expuseram a necessidade de fundamentos mais rigorosos para a matemática, a Teoria de Tipos surgiu como uma estrutura lógica para evitar inconsistências. Na computação, essa ideia se reflete nos sistemas de tipos modernos, que impedem comportamentos paradoxais ou indefinidos em programas. Por exemplo, ao restringir operações a tipos específicos, evita-se que funções sejam aplicadas a entidades incompatíveis — análogo a evitar que o barbeiro pertença ao conjunto que gera a contradição.

Atualmente, a principal aplicação da Teoria de Tipos está na formalização de sistemas de tipos para linguagens de programação. Um sistema de tipos garante a ausência de certos comportamentos dos programas classificando os valores computados em cada uma de suas sentenças (PIERCE, 2002). Além disso, atribuir e verificar tipos para cada construção presente nos programas têm várias utilidades, como fornecer informações para auxiliar na modularização de programas, otimização de código executada pelo compilador, além de poder ser usada como documentação do código.

No contexto das linguagens de programação, podemos distinguir três categorias principais de tipos: tipos simples, tipos polimórficos e tipos dependentes (PIERCE, 2002). Tipos simples atribuem um tipo fixo a cada termo, enquanto tipos polimórficos introduzem a noção de generalidade, permitindo que funções possam ser aplicadas a argumentos de diferentes tipos sem a necessidade de serem redefinidas para cada um. Já os tipos dependentes permitem que tipos dependam de valores.

Um exemplo de tipo simples é uma função que opera sobre números inteiros. Esta função recebe um número inteiro e retorna outro número inteiro. Seu tipo, portanto, é representado como $Int \rightarrow Int$, indicando que tanto a entrada quanto a saída são do tipo inteiro.

Um exemplo de polimorfismo é a função identidade, que recebe um elemento de qualquer tipo e retorna o mesmo elemento. Seu tipo é expresso como $a \rightarrow a$, onde a pode ser qualquer tipo, caracterizando o polimorfismo paramétrico. Nesse caso, a função mantém o mesmo comportamento para todos os tipos, sem necessidade de reimplementação.

Em linguagens com suporte a tipos dependentes, um exemplo seria o de um vetor cujo comprimento (número de elementos) faz parte de seu tipo. Nesse caso, uma função de concatenação de vetores deve garantir que somente vetores com tipos compatíveis em relação ao comprimento possam ser concatenados. O tipo da função de concatenação seria algo como¹ $Vector(n) \rightarrow Vector(m) \rightarrow Vector(n + m)$, onde n e m são valores que representam os comprimentos dos vetores e fazem parte da definição de tipo.

No contexto do polimorfismo, Pierce (2002) define duas principais variedades: o polimorfismo paramétrico, (como é o caso da função identidade), e o polimorfismo com sobrecarga. No primeiro, uma única definição opera genericamente, mantendo o mesmo comportamento para todos os tipos. Já no segundo, o comportamento varia conforme o tipo dos argumentos, permitindo múltiplas implementações — como na sobrecarga de operadores, onde a função

¹ A notação exata pode variar entre diferentes linguagens de programação que suportam tipos dependentes. A estrutura apresentada serve apenas como uma ilustração conceitual do comportamento esperado.

selecionada depende dos tipos dos operandos.

Em linguagens como Haskell, o sistema de tipos deduz automaticamente tipos genéricos sempre que possível, permitindo que funções como a identidade ($\lambda x.x$, tipada como $\forall \alpha. \alpha \rightarrow \alpha$) sejam usadas de forma universal. Por outro lado, funções com restrições específicas ilustram o polimorfismo com sobrecarga. Por exemplo, a função de soma no Código 2.3, originalmente definida como $Int \rightarrow Int \rightarrow Int$, pode ser generalizada para operar sobre quaisquer tipos numéricos utilizando classes de tipos. Isso é alcançado ao substituir a tipagem explícita por uma restrição como $Num\ a \Rightarrow a \rightarrow a \rightarrow a$, onde $Num\ a$ indica que a deve pertencer à classe de tipos numéricos.

Em Haskell, classes de tipos são mecanismos que habilitam o polimorfismo de sobrecarga. A função `sumList`, por exemplo, (Código 2.7) é definida com a restrição $Num\ a$, permitindo que opere sobre listas de inteiros, `Float`, `Double` ou outros tipos numéricos. Essa abordagem combina flexibilidade e segurança: o polimorfismo com sobrecarga garante que a função generalize seu comportamento apenas dentro de um domínio específico (*e.g.*, números), evitando inconsistências.

```

1 sumList :: (Num a) => [a] -> a
2 sumList [] = 0
3 sumList (x : xs) = x + sumList xs

```

Código 2.7 – Função somatório de elementos de lista em Haskell

2.3 CÁLCULO LAMBDA SIMPLESMENTE TIPADO

O Cálculo Lambda Simplesmente Tipado é uma das primeiras e mais simples variantes do Cálculo Lambda que incorpora tipos em sua estrutura (CHURCH, 1940). Enquanto o cálculo lambda original não faz distinção entre diferentes tipos de dados, no Cálculo Lambda Simplesmente Tipado os termos são anotados com tipos. Cada função recebe e retorna valores de tipos específicos, o que permite prevenir uma série de erros comuns em programas, como a aplicação de funções a argumentos incorretos. Além disso, o sistema de tipos serve como uma ferramenta de verificação durante a compilação de programas, assegurando que erros de tipo sejam detectados antes da execução. Dessa forma, ele não apenas facilita a criação de software mais robusto, mas também oferece uma base formal para o estudo de linguagens de programação (PIERCE, 2002).

A sintaxe básica do Cálculo Lambda Simplesmente Tipado inclui:

Variáveis: x, y, z, \dots
 Tipos: $\tau ::= \text{int} \mid \text{bool} \mid \tau \rightarrow \tau$
 Termos: $e ::= \lambda x: \tau. e \mid e_1\ e_2 \mid x$

No Cálculo Lambda Simplesmente Tipado, cada variável possui um tipo atribuído, e os termos são construídos com base nesses tipos. Por exemplo, a abstração de função $\lambda x: \tau. e$ define uma

função onde a variável x é de tipo τ e o corpo da função é a expressão e . A aplicação de função $e_1 e_2$ indica que a expressão e_1 é uma função aplicada ao argumento e_2 , o qual deve ter um tipo compatível com o esperado por e_1 . Essa formalização facilita a composição de funções e o raciocínio sobre a estrutura dos programas, pois cada termo pode ser avaliado dentro de um contexto de tipagem. A sintaxe dos tipos, como $\tau \rightarrow \tau$, define uma função que aceita um argumento de tipo τ e retorna um valor também do tipo τ .

A inferência de tipos no Cálculo Lambda Simplesmente Tipado assegura que cada expressão tenha um tipo bem definido, com base nas regras de tipagem. A tipagem de termos é feita por meio de um conjunto de regras formais que garantem a consistência dos tipos no programa. Por exemplo, a regra de tipagem para abstrações lambda é a seguinte:

$$\frac{\Gamma, x: \tau_1 \vdash e: \tau_2}{\Gamma \vdash (\lambda x: \tau_1. e): \tau_1 \rightarrow \tau_2}$$

Isso significa que, se a expressão e possui o tipo τ_2 sob o contexto em que x possui o tipo τ_1 , então a abstração $\lambda x: \tau_1. e$ tem o tipo $\tau_1 \rightarrow \tau_2$. Essa verificação de tipo garante que, ao aplicar a função, o tipo do argumento corresponde ao tipo esperado pela função.

O Cálculo Lambda Simplesmente Tipado está intimamente relacionado à lógica intuicionista proposicional. Esse vínculo é formalizado pela Correspondência Curry-Howard, que estabelece uma correspondência direta entre proposições lógicas e tipos, e entre provas e programas (PIERCE, 2002). Em outras palavras, tipos podem ser interpretados como proposições lógicas, e termos tipados como provas dessas proposições. Por exemplo, o tipo $A \rightarrow B$ no Cálculo Lambda Simplesmente Tipado pode ser visto como a implicação lógica “se A , então B ”. Assim, uma função que aceita um argumento do tipo A e retorna um valor do tipo B é equivalente a uma prova de que A implica B . Esse princípio permite usar ferramentas da teoria de tipos para construir provas formais de teoremas em lógica intuicionista, fornecendo uma base teórica robusta para assistentes de prova automatizados, como o Coq (COQUAND; HUET, 1988).

Além disso, a Correspondência Curry-Howard não apenas conecta tipos e lógica, mas também oferece um método sistemático para projetar e raciocinar sobre sistemas de inferência de tipos, garantindo que programas tipados sejam corretos em relação às especificações lógicas. A inferência de tipos desempenha um papel fundamental na programação funcional moderna, sendo inicialmente introduzida com a linguagem ML por Damas e Milner (1982), com o algoritmo W. A linguagem Haskell estende o sistema Damas-Milner, adicionando principalmente o suporte a sobrecarga de funções.

2.4 SISTEMA DAMAS-MILNER

O sistema Damas-Milner, introduzido por Robin Milner e posteriormente formalizado em maior detalhe por Luis Damas (MILNER, 1978; DAMAS; MILNER, 1982), é um dos sistemas de tipos mais influentes para linguagens funcionais. Este sistema tem como principal característica a inferência automática de tipos polimórficos, sem a necessidade de anotações

explícitas por parte do programador, ocorrendo em linguagens como ML, Haskell e OCaml. A sua base é o cálculo lambda com polimorfismo paramétrico, introduzido via `let`, permitindo que funções possam operar sobre múltiplos tipos de maneira genérica.

A introdução do sistema Damas-Milner trouxe duas contribuições principais: a definição de um sistema de tipos robusto e a criação de um algoritmo, o Algoritmo W, capaz de inferir o tipo mais geral (do inglês *principal type-scheme*), conforme demonstrado em Damas (1984). O algoritmo é consistente e completo em relação ao sistema de tipos: a consistência assegura que todo tipo inferido é correto, ou seja, pode ser derivado pelo sistema de tipos; já a completude garante que qualquer tipo derivado pelo sistema será uma instância do tipo inferido pelo algoritmo. Como resultado, a linguagem ML e suas derivadas se tornaram notórias por fornecer ao programador a capacidade de escrever programas sem erros de tipo detectáveis durante a compilação, permitindo um desenvolvimento mais seguro e robusto (MILNER, 1978; DAMAS, 1984).

A sintaxe do sistema Damas-Milner define as expressões e os tipos usados no processo de inferência. Abaixo, segue a gramática das expressões e tipos:

Variáveis	x
Expressões	$e ::= x \mid e' \mid \lambda x. e \mid \text{let } x = e \text{ in } e'$
Variáveis de tipo	α
Tipos primitivos	ι
Tipos	$\tau ::= \alpha \mid \iota \mid \tau \rightarrow \tau'$
Esquemas de tipo	$\sigma ::= \forall \alpha. \sigma \mid \tau$

Na sintaxe, x representa variáveis que podem ser nomes de qualquer identificador, e e descreve expressões que podem ser variáveis, aplicações de função, funções anônimas ou declarações `let`, que introduzem polimorfismo através de generalização de tipos. α é usado para representar variáveis de tipos. Os tipos primitivos ι são usados para representar tipos constantes. Tipos τ podem ser tanto variáveis de tipo quanto funções entre tipos. Por fim, σ denota os esquemas de tipo (do inglês *schemes*), ou tipos polimórficos, que podem quantificar variáveis de tipo, permitindo reutilização de variáveis de tipos em diferentes contextos.

O polimorfismo no sistema Damas-Milner é introduzido pelas expressões `let`, que permitem a generalização de tipos. Ao declarar uma variável ou função usando `let`, o tipo inferido é generalizado para ser utilizado de maneira polimórfica na expressão que ocorre após o `in`. Isso significa que, ao declarar uma função como `let id = $\lambda x. x$` , o sistema deduz o tipo mais geral $\forall \alpha. \alpha \rightarrow \alpha$, que pode ter sua variável de tipo α instanciada para diferentes tipos conforme for necessário.

A inferência de tipos envolve dois processos principais: generalização e instanciação. A generalização ocorre quando o sistema identifica que uma expressão pode ser tipada com um tipo mais geral, permitindo que seja reutilizada de maneira polimórfica. Já a instanciação

ocorre quando um tipo polimórfico é aplicado a um tipo concreto, especializando-o para um uso específico. Esse mecanismo garante a flexibilidade do sistema, ao mesmo tempo que mantém a segurança garantida pela inferência de tipos. Por exemplo, considere a expressão `let id = $\lambda x.x$ in (id 1, id 'a')`. O sistema generaliza o tipo de `id` para $\forall \alpha. \alpha \rightarrow \alpha$, e instancia este tipo tanto para inteiros quanto para caracteres nas duas aplicações subsequentes.

Outro conceito importante para o processo de inferência de tipos no sistema é a substituição de tipos, onde estes são mapeados para outros tipos ou para variáveis de tipo. Formalmente, uma substituição de tipos é representada como um mapeamento finito de variáveis de tipo para tipos, denotado por S , e pode ser escrito na forma $[\alpha_1 \mapsto \tau_1, \alpha_2 \mapsto \tau_2, \dots, \alpha_n \mapsto \tau_n]$. Aqui, α_i são variáveis de tipo distintas e τ_i são os tipos correspondentes. Em outras palavras, S associa cada variável de tipo α_i a um tipo τ_i específico.

A aplicação de uma substituição S em um tipo τ , denotada por $S\tau$, resulta na substituição de todas as ocorrências livres de α_i em τ por τ_i . Esse conceito de substituição é fundamental para o processo de instanciação de tipos, que será discutido a seguir. A definição formal da aplicação de substituições é dada por:

$$\begin{aligned} S\alpha_i &\equiv \tau_i, \\ S\alpha &\equiv \alpha, \quad \text{se } \alpha \notin \{\alpha_1, \alpha_2, \dots, \alpha_n\}, \\ S(\tau_1 \rightarrow \tau_2) &\equiv S\tau_1 \rightarrow S\tau_2, \\ S(\forall \alpha. \sigma) &\equiv S'\sigma, \quad \text{onde } S' = S \setminus [\alpha \mapsto _]. \end{aligned}$$

Figura 2.2 – Regras de aplicação da substituição de tipos

Fonte: (SILVA, 2019)

onde o símbolo de subtração de conjuntos (\setminus) indica que a substituição S' é a substituição S restrita ao conjunto de mapeamentos que não envolvem a variável α .

A instanciação de tipos é um processo em que um esquema de tipo $\sigma = \forall \alpha_1 \dots \alpha_m. \tau$ é transformado em um tipo específico substituindo suas variáveis quantificadas por tipos concretos. Se S é uma substituição, então $S\sigma$ é o esquema de tipo obtido substituindo cada ocorrência livre de α_i em σ por τ_i , renomeando as variáveis genéricas de σ , se necessário. O tipo resultante $S\sigma$ é chamado de uma instância de σ (DAMAS; MILNER, 1982). Esse processo é essencial para adaptar esquemas de tipos polimórficos a situações específicas em um programa, mantendo a flexibilidade e segurança do sistema de tipos.

Um esquema de tipo também pode ter uma instância genérica $\sigma' = \forall \beta_1 \dots \beta_n. \tau'$, se existir uma substituição $[\tau_i / \alpha_i]$ tal que $\tau' = [\tau_i / \alpha_i] \tau$, e as variáveis β_j não aparecem livres em σ . Nesse caso, escrevemos $\sigma > \sigma'$, indicando que σ é mais geral do que σ' . Vale notar que a instanciação atua sobre variáveis livres, enquanto a instanciação genérica lida com variáveis ligadas.

O sistema de tipos de Damas e Milner é definido por um conjunto de regras de inferência de tipos, apresentadas na Figura 2.3, que são usadas para determinar os tipos das expressões no

sistema. Essas regras são representadas por meio de julgamentos de tipos da forma $\Gamma \vdash e: \sigma$, onde Γ é o contexto – um conjunto de suposições na forma de pares (x_i, σ_i) , associando variáveis x_i aos seus respectivos tipos σ_i –, e é a expressão sendo tipada, e σ é o tipo inferido para essa expressão.

$$\boxed{\Gamma \vdash x: \tau}$$

$$\begin{aligned} [\text{Taut}]: \quad & \frac{x: \sigma \in \Gamma}{\Gamma \vdash x: \sigma} \\ [\text{Abs}]: \quad & \frac{\Gamma, x: \tau \vdash e: \tau'}{\Gamma \vdash (\lambda x. e): \tau \rightarrow \tau'} \\ [\text{App}]: \quad & \frac{\Gamma \vdash e: \tau' \rightarrow \tau \quad \Gamma \vdash e': \tau'}{\Gamma \vdash (e \ e'): \tau} \\ [\text{Let}]: \quad & \frac{\Gamma \vdash e: \sigma \quad \Gamma, x: \sigma \vdash e': \tau}{\Gamma \vdash (\text{let } x = e \text{ in } e'): \tau} \\ [\text{Inst}]: \quad & \frac{\Gamma \vdash e: \sigma}{\Gamma \vdash e: \sigma'} \ (\sigma > \sigma') \\ [\text{Gen}]: \quad & \frac{\Gamma \vdash e: \sigma}{\Gamma \vdash e: \forall \alpha. \sigma} \ \alpha \notin \text{FV}(\Gamma) \end{aligned}$$

Figura 2.3 – Regras de Inferência do sistema Damas-Milner
Fonte: o autor. Adaptado de (DAMAS; MILNER, 1982)

As regras de inferência são interpretadas de baixo para cima. Por exemplo, na regra da tautologia [Taut] apresentada na Figura 2.3, significa que, se em um contexto Γ , a variável x possui o tipo σ , então podemos concluir que x tem o tipo σ no mesmo contexto. Isso reflete o fato de que a associação de tipos no contexto é preservada.

Na regra de generalização [Gen] da Figura 2.3, a condição de que α não seja livre em Γ — formalmente escrita como $\alpha \notin \text{FV}(\Gamma)$ — assegura que o tipo generalizado não dependa de nenhum tipo específico presente no contexto. Isso permite que o tipo $\forall \alpha. \sigma$ seja usado de forma polimórfica em diferentes partes do programa.

Essas regras garantem a solidez do sistema, preservando a segurança dos tipos ao inferir automaticamente os tipos mais gerais possíveis para as expressões.

Antes de apresentar o Algoritmo W, é importante observar que ele é uma implementação prática das regras de inferência aqui descritas, usando o conceito de unificação para resolver as equações de tipo geradas durante a inferência. A seguir, será discutido em detalhes o funcionamento do Algoritmo W.

2.4.1 Algoritmo W

O Algoritmo W, introduzido em Damas e Milner (1982), é um algoritmo eficiente² para inferência de tipos em linguagens de programação funcional. Ele se baseia no processo de unificação para resolver equações de tipos geradas durante a análise de expressões, atribuindo a cada expressão um tipo principal (do inglês *principal type-scheme*) (isto é, o tipo mais geral possível, no sentido de que qualquer outro tipo atribuível à expressão pode ser obtido a partir deste por substituição). Formalmente, para uma expressão e , o Algoritmo W encontra um tipo τ tal que, para qualquer tipo τ' que também possa ser atribuído a e , existe uma substituição S tal que $S(\tau) = \tau'$.

A unificação é o processo de encontrar uma substituição de variáveis de tipo que torna dois tipos dados equivalentes. Formalmente, dados dois tipos τ_1 e τ_2 , a unificação procura uma substituição S tal que $S\tau_1 = S\tau_2$. Se tal substituição existe, os tipos são considerados unificáveis e S é chamada de solução unificadora. Caso contrário, os tipos são incompatíveis.

O algoritmo de unificação, *unify*, descrito na Figura 2.4, retorna a substituição que representa o unificador mais geral, operando recursivamente sobre a estrutura dos tipos. Ele verifica se os tipos são idênticos, se uma variável de tipo pode ser substituída por outro tipo, ou, no caso de tipos compostos, se suas partes podem ser unificadas independentemente.

```

unify( $\alpha$ ,  $\alpha$ ) =
  retorna [ ]
unify( $\alpha$ ,  $\tau$ ) =
  se occurs( $\alpha$ ,  $\tau$ ), então
    falha
  senão
    retorna [ $\alpha \mapsto \tau$ ]
unify( $\tau$ ,  $\alpha$ ) =
  se occurs( $\alpha$ ,  $\tau$ ), então
    falha
  senão
    retorna [ $\alpha \mapsto \tau$ ]
unify( $\tau_1 \rightarrow \tau_2$ ,  $\tau'_1 \rightarrow \tau'_2$ ) =
  retorna unify( $\tau_1$ ,  $\tau'_1$ )  $\circ$  unify( $\tau_2$ ,  $\tau'_2$ )

```

Figura 2.4 – Algoritmo de unificação para o Sistema Damas-Milner no formato de função.

Fonte: autor. Adaptado de (RIBEIRO; CAMARÃO, 2016)

O algoritmo *unify* faz uso da função de verificação de ocorrência *occurs* apresentado na Figura 2.5, que por sua vez, tem como propósito evitar substituições que introduzam ciclos, como

² Embora seja eficiente na grande maioria dos casos, há situações em que o Algoritmo W apresenta desempenho exponencial, conforme discutido em (VASCONCELLOS, 2004).

$[\alpha \mapsto \alpha \rightarrow \alpha]$, que resultaria em inconsistência no sistema de tipos (RIBEIRO; CAMARÃO, 2016). Esta função verifica recursivamente se uma variável de tipo α aparece em um tipo τ . Na última regra do algoritmo *unify*, ao unificar dois tipos função (do inglês *arrow type*) $\tau_1 \rightarrow \tau_2$ e $\tau'_1 \rightarrow \tau'_2$, a substituição resultante é obtida pela composição $S_2 \circ S_1$, onde $S_1 = \text{unify}(\tau_1, \tau'_1)$ e $S_2 = \text{unify}(\tau_2, \tau'_2)$. Aqui, a composição $S_2 \circ S_1$ é definida como a substituição que aplica S_1 primeiro, seguida por S_2 , ou seja, para qualquer variável de tipo α , temos $(S_2 \circ S_1)(\alpha) = S_2(S_1(\alpha))$. Essa definição garante que as substituições sejam aplicadas corretamente em cascata durante o processo de unificação.

```

occurs( $\alpha$ ,  $\tau_1 \rightarrow \tau_2$ ) =
    retorna occurs( $\alpha$ ,  $\tau_1$ )  $\vee$  occurs( $\alpha$ ,  $\tau_2$ )
occurs( $\alpha$ ,  $\alpha$ ) =
    retorna Verdadeiro
occurs( $\alpha$ ,  $\tau$ ) =
    retorna Falso

```

Figura 2.5 – Algoritmo de verificação de ocorrência para o Sistema Damas-Milner no formato de função.

Fonte: autor. Adaptado de (RIBEIRO; CAMARÃO, 2016)

O Algoritmo W, descrito na Figura 2.6 é um método utilizado para inferência de tipos em expressões de linguagens funcionais. Ele atribui os tipos mais gerais possíveis a cada subexpressão, garantindo a consistência com as operações definidas. O algoritmo combina a unificação com regras de inferência de tipos para deduzir o tipo de uma expressão, explorando o polimorfismo de forma eficiente.

$\Gamma \vdash_W e: \tau, S$

$$\begin{aligned}
 &[\text{Var}] \frac{x: \sigma \in \Gamma \quad \tau = \text{inst}(\sigma)}{\Gamma \vdash_W x: \tau, \emptyset} \\
 &[\text{App}] \frac{\Gamma \vdash_W e_0: \tau_0, S_0 \quad S_0 \Gamma \vdash_W e_1: \tau_1, S_1 \quad \tau' = \text{newvar} \quad S_2 = \text{mgu}(S_1 \tau_0, \tau_1 \rightarrow \tau')}{\Gamma \vdash_W e_0 e_1: S_2 \tau', S_2 \circ S_1 \circ S_0} \\
 &[\text{Abs}] \frac{\tau = \text{newvar} \quad \Gamma, x: \tau \vdash_W e: \tau', S}{\Gamma \vdash_W \lambda x. e: S \tau \rightarrow \tau', S} \\
 &[\text{Let}] \frac{\Gamma \vdash_W e_0: \tau, S_0 \quad S_0 \Gamma, x: \overline{S_0 \Gamma}(\tau) \vdash_W e_1: \tau', S_1}{\Gamma \vdash_W \text{let } x = e_0 \text{ in } e_1: \tau', S_1 \circ S_0}
 \end{aligned}$$

Figura 2.6 – Algoritmo W em formato de regras de inferência.

Fonte: o autor. Adaptado de (SILVA, 2019)

O funcionamento pode ser analisado para os diferentes tipos de expressões a seguir. Para uma variável x (regra [Var]), o algoritmo verifica se existe um esquema de tipo σ associado a x no contexto Γ . Quando presente, realiza-se a instanciação de σ substituindo suas variáveis quantificadas por tipos frescos (do inglês *fresh types*) (variáveis novas sem colisões), obtendo o tipo concreto τ . A substituição identidade (\emptyset) é retornada com o tipo inferido.

Para aplicações de função $e_0 e_1$ (regra [App]), o algoritmo primeiro infere recursivamente o tipo de e_0 em Γ , obtendo τ_0 e a substituição S_0 . Em seguida, no contexto atualizado $S_0\Gamma$, infere o tipo de e_1 obtendo τ_1 e S_1 . Um novo tipo variável fresco τ' é introduzido, e o unificador mais geral (mgu) reconcilia $S_1\tau_0$ com $\tau_1 \rightarrow \tau'$, produzindo a substituição S_2 . O tipo resultante $S_2\tau'$ e a composição de substituições $S_2 \circ S_1 \circ S_0$ são retornados.

No caso de abstrações $\lambda x.e$ (regra [Abs]), um tipo fresco τ é gerado para x . O contexto é estendido para $\Gamma, x: \tau$, e o tipo de e é inferido nesse novo contexto, produzindo τ' e a substituição S . O tipo da abstração é então determinado como $S\tau \rightarrow \tau'$, mantendo-se a substituição S .

Para expressões $\text{let } x = e_0 \text{ in } e_1$ (regra [Let]), o tipo de e_0 é inferido primeiro, resultando em τ e S_0 . O tipo τ é generalizado no contexto $S_0\Gamma$ através da operação $\overline{S_0\Gamma}(\tau)$, definida formalmente como:

$$\overline{S_0\Gamma}(\tau) = \forall \hat{\alpha}. \tau \quad \text{onde} \quad \hat{\alpha} = \text{free}(\tau) - \text{free}(S_0\Gamma)$$

ou seja, as variáveis de tipo $\hat{\alpha}$ (que são livres em τ mas não no contexto $S_0\Gamma$) são quantificadas, gerando um esquema de tipo polimórfico. Este esquema é associado a x ao inferir o tipo de e_1 no contexto $S_0\Gamma$, obtendo τ' e S_1 . O tipo final τ' e a composição $S_1 \circ S_0$ são retornados.

2.5 CÁLCULO DE CPS SIMPLEMENTE TIPADO

O Cálculo de CPS se origina como a IR do compilador descrita no livro do Appel (1992). Sua versão simplesmente tipada foi estudada na tese do Thielecke (1997), tendo sido descrita de maneira categórica. Neste trabalho entretanto, é adotada uma versão baseada em Teoria de Tipos, como apresentada em Torrens, Orchard e Vasconcellos (2024).

O cálculo de continuções (do inglês, *CPS-calculus*), é um sistema formal que leva o CPS além de seu uso tradicional como uma técnica de transformação de código, tratando-o como um modelo computacional por si só. Enquanto o CPS é utilizado como uma IR em compiladores, o cálculo de continuções oferece uma estrutura para raciocinar formalmente sobre computações onde o fluxo de controle é explicitamente representado. A sintaxe deste cálculo é descrita a seguir:

$$\text{Comandos } b ::= x(\vec{x}) \mid \text{let } x(\vec{x}) = b \text{ in } b$$

Vale ressaltar que Thielecke (1997) possui uma sintaxe diferente para o cálculo de continuções, onde os termos são respectivamente representados como $k\langle\vec{x}\rangle$ e $b \{ x\langle\vec{x}\rangle = b \}$. Em razão da natureza prática e do público alvo deste trabalho, a notação adotada para o *CPS-Calculus* será a

do Appel. Aqui, $k(\vec{x})$ representa um salto (do inglês *jump*), isto é, uma chamada direta para a continuação k com os parâmetros \vec{x} , essencialmente passando o controle para essa k . Enquanto que $\text{let } k(\vec{y}) = c \text{ in } b$ representa um vínculo (do inglês *binding*), onde a continuação k com os parâmetros \vec{y} é ligada dentro do comando b com o corpo c . Isto modela uma construção intermediária em que se define uma nova continuação que poderá ser chamada dentro de b para prosseguir com a computação.

Neste sistema, há apenas um construtor de tipos: a negação poliádica (do inglês *polyadic negation type*), usada para representar continuações (TORRENS; ORCHARD; VASCONCELLOS, 2024). Além disso, X varia sobre os tipos base.

$$\begin{array}{ll} \text{Tipos} & \tau ::= \neg \vec{\tau} \mid X \\ \text{Contexto} & \Gamma ::= \cdot \mid \Gamma, x: \tau \end{array}$$

O sistema apresenta somente duas regras de tipagem principais, uma para o *jump* e outra para o *bind*.

$$\boxed{\Gamma \vdash b}$$

$$\frac{\Gamma(k) = \neg \vec{\tau} \quad \Gamma(\vec{x}) = \vec{\tau}}{\Gamma \vdash k(\vec{x})} [\text{J}]$$

$$\frac{\Gamma, k: \neg \vec{\tau} \vdash b \quad \Gamma, \vec{x}: \vec{\tau} \vdash c}{\Gamma \vdash \text{let } k(\vec{x}) = c \text{ in } b} [\text{B}]$$

Os julgamentos de tipo nesse cálculo são unilaterais, no sentido de que tipos aparecem apenas em um dos lados do *turnstile* (\vdash). Diferentemente do cálculo de tipos tradicional como o de Damas-Milner, onde um julgamento do tipo $\Gamma \vdash e: \tau$ atribui um tipo τ à expressão e sob o contexto Γ , no cálculo de CPS é escrito apenas $\Gamma \vdash b$. Isso significa que, sob o contexto Γ , o termo b é consistente, ou seja, não vai falhar durante a execução (THIELECKE, 1997). Por exemplo, a derivação $x: \tau, k: \neg \tau \vdash k(x)$ indica que, assumindo x com tipo τ e k com tipo $\neg \tau$, passar x como argumento para k é seguro.

Torrens, Orchard e Vasconcellos (2024) demonstram a segurança de tipos (do inglês *type safety*) deste sistema, provando as propriedades de preservação e progresso. Isso garante que expressões bem tipadas não entram em estados inválidos ao serem avaliadas.

Um contraste interessante pode ser observado entre o sistema Damas-Milner e o cálculo de CPS. No primeiro, as expressões são testemunhas da veracidade de seus tipos. Já no segundo, as expressões podem ser vistas como testemunhas de que o contexto de tipos leva a uma contradição — nas palavras de Thielecke, “ b witnesses that Γ entails a contradiction”. Como no exemplo do $x: \tau, k: \neg \tau \vdash k(x)$, observa-se que $k(x)$ representa uma contradição, ao termos tanto τ (em x) como $\neg \tau$ (em k).

Enquanto o cálculo- λ define sua semântica por meio das regras β e η , o *CPS-calculus* possui uma teoria equacional própria, formulada em Thielecke (1997). Esta teoria é composta

por quatro axiomas fundamentais que permitem raciocinar sobre a equivalência de programas escritos nessa representação. Esses axiomas são:

$$(JMP) \quad k\langle \vec{x} \rangle \{ k\langle \vec{y} \rangle = c \} = c[\vec{x}/\vec{y}] \quad \text{se } k \notin \vec{x}$$

O *Jump (JMP)* realiza uma chamada direta à continuação k , substituindo os parâmetros formais \vec{y} pelos argumentos reais \vec{x} no corpo c . A necessidade da condição $k \notin \vec{x}$ é tal que esta evita capturas.

$$(GC) \quad k\langle \vec{x} \rangle \{ j\langle \vec{y} \rangle = c \} = k\langle \vec{x} \rangle \quad \text{se } j \notin FV(k\langle \vec{x} \rangle)$$

O axioma *Garbage Collection (GC)*, elimina os vínculos de continuações não referenciadas em c .

$$(ETA) \quad b \{ k\langle \vec{x} \rangle = j\langle \vec{x} \rangle \} = b[j/k] \quad \text{se } j \notin \vec{x}$$

O axioma da extensionalidade (ETA) diz respeito as continuações redundantes, onde k é definido apenas como uma chamada direta a j .

$$(DISTR) \quad b \{ k\langle \vec{x} \rangle = c \} \{ j\langle \vec{y} \rangle = d \} = b \{ j\langle \vec{y} \rangle = d \} \{ k\langle \vec{x} \rangle = c \} \{ j\langle \vec{y} \rangle = d \} \\ \text{se } k \neq j, j \notin \vec{x}, \{ k, \vec{x} \} \notin FV(d)$$

Por fim, a distribuição, com o axioma (DISTR) reordena os vínculos para expor saltos.

2.5.1 Tradução CPS

O processo de tradução, no contexto de linguagens é a etapa em que um código feito em uma linguagem é transcrito para outra, onde idealmente, semântica alguma deve ser perdida. Plotkin (1975) introduziu duas traduções do cálculo- λ , posteriormente adaptadas ao cálculo de CPS por Thielecke (1997), diferenciando-se conforme a estratégia de avaliação adotada. A tradução por chamada por nome (CBN, do inglês *call-by-name*) permite simular a execução no cálculo- λ , representada por $\llbracket - \rrbracket_N$; de maneira análoga, a tradução por chamada por valor (CBV, do inglês *call-by-value*) é representada por $\llbracket - \rrbracket_V$ no cálculo- λ_V . Embora essas traduções não sejam as mais eficientes, foram escolhidas por sua clareza e simplicidade.

$$\boxed{\llbracket e \rrbracket_N = b} \text{ Plotkin's call-by-name translation}$$

$$\begin{aligned} \llbracket x \rrbracket_N &= x(k) \\ \llbracket \lambda x. e \rrbracket_N &= \text{let } v(k, x) = \llbracket e \rrbracket_N \text{ in } k(v) \\ \llbracket f e \rrbracket_N &= \text{let } k(f) = \text{let } v(k) = \llbracket e \rrbracket_N \text{ in } c \text{ in } \llbracket f \rrbracket_N \\ \llbracket \text{let } x = f \text{ in } e \rrbracket_N &= \text{let } x(k) = \llbracket f \rrbracket_N \text{ in } \llbracket e \rrbracket_N \end{aligned}$$

$$\boxed{\llbracket e \rrbracket_V = b} \text{ Plotkin's call-by-value translation}$$

$$\begin{aligned}
\llbracket x \rrbracket_V &= k(x) \\
\llbracket \lambda x. e \rrbracket_V &= \text{let } v(k, x) = \llbracket e \rrbracket_V \text{ in } k(v) \\
\llbracket f e \rrbracket_V &= \text{let } k(f) = \text{let } v(k) = c \text{ in } \llbracket e \rrbracket_V \text{ in } \llbracket f \rrbracket_V \\
\llbracket \text{let } x = f \text{ in } e \rrbracket_V &= \text{let } k(x) = \llbracket e \rrbracket_V \text{ in } \llbracket f \rrbracket_V
\end{aligned}$$

As variáveis f , v e k que não aparecem no termo de origem são consideradas frescas, onde as duas primeiras são imediatamente ligadas, enquanto que k é esperada que seja livre e por ser vinculada às traduções dos subtermos, um único k é necessário (TORRENS; ORCHARD; VASCONCELLOS, 2024).

Ainda, a tradução do ‘let’ não está presente no artigo. Esta foi feita para que o trabalho em questão pudesse ser mais completo e mais próximo de algo utilizável em um caso real de implementação de compilador com uso desta IR. A maneira de se obtê-la foi a partir de uma beta-redex, simplificando a tradução de $((\lambda x. b) a)$ por meio dos axiomas.

Além da tradução de expressões, uma vez que o ambiente sendo trabalhado é tipado, é necessário que a tradução preserve a tipagem dos programas, Torrens, Orchard e Vasconcellos (2024) apresentam a tradução tipada do CPS, onde, para o tipo funcional $A ::= A \rightarrow A \mid X$, as seguintes traduções são definidas. Note que aqui também as funções são definidas diferentemente de acordo com estratégia de avaliação adotada:

$$\begin{aligned}
\llbracket X \rrbracket_N &= \neg X \\
\llbracket A \rightarrow B \rrbracket_N &= \neg \neg (\neg \llbracket A \rrbracket_N, \llbracket B \rrbracket_N) \\
\llbracket \vec{x}: \vec{A} \vdash e: B \rrbracket_N &= \vec{x}: \neg \llbracket A \rrbracket_N, k: \llbracket B \rrbracket_N \vdash \llbracket e \rrbracket_N \\
\llbracket X \rrbracket_V &= X \\
\llbracket A \rightarrow B \rrbracket_V &= \neg (\llbracket A \rrbracket_V, \neg \llbracket B \rrbracket_V) \\
\llbracket \vec{x}: \vec{A} \vdash e: B \rrbracket_V &= \vec{x}: \llbracket A \rrbracket_V, k: \neg \llbracket B \rrbracket_V \vdash \llbracket e \rrbracket_V
\end{aligned}$$

Esta função de tradução tem como propósito, mostrar que se o termo lambda é tipado, então o termo traduzido para CPS também é tipado. É importante notar que isto se aplica somente para os tipos simples, como definido no tipo funcional a partir do tipo primitivo X e do *arrow type* $A \rightarrow A$. O sistema de tipos proposto neste trabalho, entretanto, é polimórfico, ou seja, aqui há a adição de variáveis de tipos quantificadas. Sendo assim, esta função de tradução, se aplicando neste caso de uso, não necessariamente retornará o tipo mais geral de uma expressão, mas sempre um subtipo deste.

3 DESENVOLVIMENTO

Será abordada, neste capítulo, a contribuição prática do trabalho, apresentando a formalização do sistema de tipos proposto, sua implementação na linguagem Haskell, e os resultados experimentais obtidos. No que diz respeito ao sistema de tipos, em virtude da limitação de tempo e do escopo deste trabalho, não foi feita a prova de que este é correto ou completo. Ao invés disso, há evidências empíricas, baseadas em testes, para suportar o correção do sistema de tipos proposto. Uma explicação mais detalhada sobre esta parte será dada na Seção 3.1.

Na Seção 3.2, serão descritas as decisões de projeto que orientaram a implementação. O ambiente de desenvolvimento utilizado foi o compilador *The Glorious Glasgow Haskell Compilation System* (GHC), na versão 9.12.2 em conjunto com o gerenciador de projetos Cabal na versão 3.14.1. O código-fonte completo está disponível publicamente no repositório *cps-type-inferer*¹ do GitHub. A Seção 3.2 terá um aprofundamento maior na implementação.

3.1 FORMALIZAÇÃO

O sistema de tipos formalizado aqui foi fortemente inspirado no sistema de Damas e Milner, explicado na Seção 2.4, onde suas regras foram adaptadas de modo que elas se enquadrem no sistema polimórfico baseado em continuações. Em particular, o contexto Γ associa variáveis a tipos polimórficos e define julgamentos distintos para representar átomos e comandos. A distinção destes se mostra necessária uma vez que é levado em consideração o comportamento não retornável das continuações.

A sintaxe do sistema conta com expressões e tipos usados no processo de tipagem e de inferência de tipos. Abaixo, segue a gramática das expressões e tipos presentes:

Átomos	$a ::= x \mid n$
Comandos	$b ::= x(\vec{a}) \mid \text{let } x(\vec{x}) = b \text{ in } b$
Tipos monomórficos	$\tau ::= \alpha \mid \text{int} \mid \neg \vec{\tau}$
Tipos polimórficos	$\sigma ::= \forall \vec{\alpha}. \tau$
Contexto	$\Gamma ::= \cdot \mid \Gamma, x: \sigma$

Na sintaxe apresentada, a representa os átomos. Isto é, variáveis do programa (x) e literais inteiros (n) formam os elementos primitivos do sistema. Os comandos b , por sua vez, são as expressões, explicadas com mais detalhes na Seção 2.1.1, sendo a primeira o *jump*, e a segunda o *bind*. Três elementos distintos compõem os tipos presentes neste sistema. Os tipos monomórficos (τ), são os tipos que não possuem quantificação, podendo ser variáveis de tipo (α), tipos numéricos inteiros (int), ou ainda, tipos negados ($\neg \vec{\tau}$), usados para representar funções que retornam absurdos. Já os tipos polimórficos (σ), são responsáveis por garantir a quantificação universal de variáveis de

¹ <<https://github.com/bidinpithecus/cps-type-inferer>>

tipos. Por fim, o contexto (Γ) contém o mapeamento de cada variável para um tipo polimórfico (σ).

As regras sintáticas de tipagem do sistema de tipos, inspiradas no Sistema Damas-Milner são ilustradas a seguir:

$$\boxed{\Gamma \vdash a: \tau}$$

$$\frac{x: \sigma \in \Gamma \quad \sigma > \tau}{\Gamma \vdash x: \tau} [\text{Var}]$$

$$\frac{}{\Gamma \vdash n: \text{int}} [\text{Int}]$$

$$\boxed{\Gamma \vdash b}$$

$$\frac{\Gamma \vdash k: \neg \vec{\tau} \quad \Gamma \vdash \vec{a}: \vec{\tau}}{\Gamma \vdash k(\vec{a})} [\text{Jump}]$$

$$\frac{\Gamma, \vec{x}: \vec{\tau} \vdash c \quad \Gamma, k: \bar{\Gamma}(\neg \vec{\tau}) \vdash b}{\Gamma \vdash \text{let } k(\vec{x}) = c \text{ in } b} [\text{Bind}]$$

Aqui, a regra [Var] define como tipo de uma variável uma instância do tipo (possivelmente polimórfico) que está associado a variável no contexto de tipos. O símbolo $>$ denota essa relação de ordem, indicando que o tipo σ é mais geral que τ . Assim, em um contexto Γ , uma variável x terá tipo τ caso esta esteja presente no contexto. A regra [Int], é direta. Em um contexto Γ , um literal inteiro terá um tipo int . Por exemplo, se $x: \forall \vec{\alpha}. \tau \in \Gamma$, então $\Gamma \vdash x: \tau$ (após instanciação adequada das variáveis de tipo).

As continuções, como discutido na Seção 2.1.1, representam fluxos de controle que não retornam valores. Uma vez que a continuação pode ser interpretada como o próximo passo de uma computação, e a computação se dá por contradições, a continuação em si não possui um tipo, ela representa um absurdo. Então, pode se dizer que a continuação é uma testemunha de que aquilo é um absurdo.

A regra [Jump] portanto, diz que sob um contexto Γ , se $k: \neg(\tau_1, \dots, \tau_n)$ com n argumentos e cada argumento a_i tiver um tipo correspondente τ_i , então $k(\vec{a})$ é válido, ou seja, o salto k com os argumentos \vec{a} é testemunha de uma contradição. De modo semelhante para o [Bind], as premissas para c e b onde c é derivável a partir do contexto $\Gamma \cup \{ \vec{x}: \vec{\tau} \}$, e b sob o contexto $\Gamma \cup \{ k: \bar{\Gamma}(\neg \vec{\tau}) \}$, são testemunhas de que o comando $\text{let } k(\vec{x}) = c \text{ in } b$ é uma contradição. Assim como o ‘let’ introduz o polimorfismo no sistema Damas-Milner, a generalização $\bar{\Gamma}(\neg \vec{\tau})$ presente na premissa do [Bind] quantifica as variáveis livres de $\vec{\tau}$ em Γ , estendendo o polimorfismo também ao CPS.

O algoritmo de inferência de tipos segue o mesmo esquema de Damas-Milner (algoritmo W) adaptado ao CPS. Assim como o W, este faz uso do unificador mais geral, retornando sempre

que existir o tipo mais genérico das expressões pertencentes a este sistema. Abaixo, tem-se sua definição:

$$\boxed{\Gamma \vdash_W a: \tau}$$

$$\frac{x: \sigma \in \Gamma \quad \tau = \text{inst}(\sigma)}{\Gamma \vdash_W x: \tau}$$

$$\frac{}{\Gamma \vdash_W n: \text{int}}$$

$$\boxed{\Gamma \vdash_W b \Rightarrow S}$$

$$\frac{\Gamma \vdash_W k: \tau_1 \quad \Gamma \vdash_W \vec{a}: \vec{\tau}_2 \quad S = \text{mgu}(\tau_1, \neg \vec{\tau}_2)}{\Gamma \vdash_W k(\vec{a}) \Rightarrow S}$$

$$\frac{\vec{\tau} = \overline{\text{newvar}} \quad \Gamma, \vec{x}: \vec{\tau} \vdash_W c \Rightarrow S_1 \quad \sigma = \overline{S_1 \Gamma}(S_1 \neg \vec{\tau}) \quad S_1 \Gamma, k: \sigma \vdash_W b \Rightarrow S_2}{\Gamma \vdash_W \text{let } k(\vec{x}) = c \text{ in } b \Rightarrow S_2 \circ S_1}$$

Para as regras de inferência dos átomos, tal qual o sistema de tipos definido anteriormente, o algoritmo com uma variável x de tipo polimórfico σ pertencente ao contexto Γ retornará um tipo monomórfico τ sob o mesmo contexto onde τ será a instanciação deste tipo σ . Para o tipo numérico, a regra não exige condições adicionais, e o algoritmo inferirá diretamente o tipo int para o literal n .

$$\frac{a: \forall \alpha. \alpha \in \Gamma \quad \alpha = \text{inst}(\forall \alpha. \alpha)}{\Gamma \vdash_W a: \alpha}$$

Por exemplo, esteja a variável a com tipo $\forall \alpha. \alpha$ no contexto, ou seja, $a: \forall \alpha. \alpha \in \Gamma$. O algoritmo então inferirá, que a variável a terá tipo α , após as devidas normalizações (redução- α).

Os comandos serão inferidos com substituições, onde o algoritmo as retornará representando o absurdo para qual esses testemunham. Para o [Jump], partindo das premissas onde sob um contexto Γ , a chamada k terá um tipo monomórfico τ_1 , os n argumentos em \vec{a} terão n tipos monomórficos τ_2 , e ainda, S é a unificação mais geral entre τ_1 e $\neg \vec{\tau}_2$, o algoritmo irá então retornar esta substituição S para o salto $k(\vec{a})$.

$$\frac{\Gamma \vdash_W k: \alpha \quad \Gamma \vdash_W x: \beta \quad S = \text{mgu}(\alpha, \neg \beta)}{\Gamma \vdash_W k(x) \Rightarrow \{ \alpha \mapsto \neg \beta \}}$$

Por exemplo, em determinado contexto Γ , seja a chamada k com tipo α , ou seja, $\Gamma \vdash_W k: \alpha$, e ainda sob o mesmo contexto, o argumento x com tipo β , ou seja, $\Gamma \vdash_W x: \beta$. A partir da unificação mais geral entre α e $\neg \beta$ é obtida a substituição S , ou seja, $S = \text{mgu}(\alpha, \neg \beta)$. O algoritmo então, irá inferir que a substituição para que o salto represente uma contradição é $\{ \alpha \mapsto \neg \beta \}$.

Vale destacar que o algoritmo de unificação apresentado na Figura 2.4, ainda que retorne a substituição que representa a unificação mais geral, não é o mesmo que o mgu utilizado na

regra [Jump]. Suas definições variam conforme os sistemas de tipos que elas atendem. Enquanto que a função *unify* é definida para os tipos do Sistema Damas-Milner, a *mgu* apresentada na Figura 3.1 é definida para os tipos do Sistema baseado em continuções.

```

mgu( $\alpha$ ,  $\tau$ ) =
    retorna varBind( $\alpha$ ,  $\tau$ )
mgu( $\tau$ ,  $\alpha$ ) =
    retorna varBind( $\alpha$ ,  $\tau$ )
mgu(Int, Int) =
    retorna [ ]
mgu(Neg list1, Neg list2) =
    se length list1  $\neq$  length list2, então falha
    senão retorna mguList(list1, list2)
mgu( $\tau_1$ ,  $\tau_2$ ) =
    falha

mguList([ ], [ ]) =
    retorna [ ]
mguList( $[\tau : \tau_s]$ ,  $[\tau' : \tau'_s]$ ) =
     $S_1 := \textit{mgu}(\tau, \tau')$ 
     $S_2 := \textit{mguList}(S_1(\tau_s), S_1(\tau'_s))$ 
    retorna  $S_2 \circ S_1$ 
mguList(_, _) =
    falha

```

Figura 3.1 – Algoritmo de unificação para o Sistema baseado em continuções no formato de função.

Fonte: autor.

A principal diferença deste algoritmo em relação ao anterior é que, neste sistema de tipos, há listas de tipos que também precisam ser unificadas. Para realizar essa unificação, verifica-se inicialmente se as duas listas possuem o mesmo tamanho, ou seja, se *length list*₁ = *length list*₂. Caso essa condição seja satisfeita, o algoritmo *mgu* é aplicado recursivamente a cada par de tipos correspondentes nos mesmos índices das listas, acumulando as substituições parciais ao longo do processo. A função *varBind* é responsável por realizar a unificação entre variáveis de tipo e outros tipos, retornando a substituição correspondente ou um erro, caso a verificação de ocorrência via *occurs* falhe.

Para a regra que garante o polimorfismo do sistema, o [Bind], a chamada *k* recebe \vec{x} argumentos, onde estes terão $\vec{\tau}$ tipos como sendo variáveis de tipo, ou seja, $\vec{\tau} = \overrightarrow{\textit{newvar}}$. O *c*, por se tratar de um comando, é uma substituição S_1 , onde recursivamente será inferida com o

contexto inicial Γ unido com os \vec{x} argumentos tipados com suas $\vec{\tau}$ variáveis de tipo frescas, ou seja, $\{ \Gamma \cup \{ \vec{x}: \vec{\tau} \} \} \vdash_W c \Rightarrow S_1$. Um ponto de atenção é necessário na função de generalização $\sigma = \overline{S_1 \Gamma}(\overline{S_1 \neg \vec{\tau}})$. A substituição S_1 aplicada no contexto garante que este esteja atualizado com a descoberta do comando c na premissa anterior. Como as continuções não retornam e simplesmente passam o resultado da computação adiante, é necessário também que S_1 seja aplicado no tipo do argumento $S_1 \neg \vec{\tau}$, para garantir que a substituição obtida no comando anterior seja utilizada nos tipos. De maneira semelhante ao primeiro comando, o comando b é inferido recursivamente com S_1 aplicado no contexto unido com o salto k tendo o tipo polimórfico σ produzido na premissa anterior, sendo atribuído a esta inferência a substituição S_2 , ou seja, $\{ S_1 \Gamma \cup \{ k: \sigma \} \} \Rightarrow S_2$. O algoritmo portanto, para o comando `let $k(\vec{x}) = c$ in b` , irá produzir a substituição resultante da composição entre as substituições de b e c , ou seja, $S_2 \circ S_1$.

3.2 IMPLEMENTAÇÃO

Partindo para a parte prática do trabalho, os módulos e funções serão apresentados de maneira gradual, de modo a facilitar o entendimento do fluxo inteiro do programa. Inicialmente, na Subseção 3.2.1, será discutido sobre a maneira como foi representado o sistema de tipos. Em sequência, a Subseção 3.2.2 trará detalhes sobre a implementação das funções de tradução de cálculo lambda simplesmente tipado para CPS. Posteriormente, a Subseção 3.2.3 irá tratar da inferência em si, juntamente da verificação do tipo. Por fim, a geração de código abordada na Subseção 3.2.4 serve como uma maneira de se testar a tradução do código.

3.2.1 Tipos de Dados

Para representar os comandos, bem como os tipos do sistema, foram utilizados tipos de dados algébricos (ADTs, do inglês *algebraic data types*), disponíveis no Código 3.1.

```

1 type Id = String
2
3 data Command
4   = Jump Id [Id]
5   | Bind Command Id [Id] Command
6
7 data CPSMonoType
8   = TVar Id
9   | TInt
10  | TNeg [CPSMonoType]
11
12 data CPSPolyType
13   = Forall [Id] CPSMonoType
14
15 type Context = Data.Map Id CPSPolyType

```



```
16 type Substitution = Data.Map Id CPSMonoType
```

Código 3.1 – Definição dos tipos de dados

Para os comandos, dois construtores podem ser observados, o *Jump* e o *Bind*, sendo responsáveis por construir respectivamente os comandos de *Jump*, onde há um salto *Id* com *[Id]* parâmetros. Sendo assim, o salto $k(x)$ seria representado por este ADT: *Jump* *k* *[x]*. E o comando *Bind*, onde há outro *Command* definido recursivamente, a função *Id* com argumentos *[Id]* e por fim outro *Command* também definido recursivamente. O *bind* *let* $k(x) = k(x)$ *in* $k(x)$, portanto seria definido pelo seguinte ADT: *Bind* (*Jump* *k* *[x]*) *k* *[x]* (*Jump* *k* *[x]*).

Os tipos foram representados com dois tipos algébricos diferentes, um para os tipos monomórficos *CPSMonoType*, e um para os tipos polimórficos *CPSPolyType*. As variáveis do tipo monomórfico são construídas a partir dos construtores *TVar Id* e *TInt*, onde no primeiro, o *Id* serve para obter a variável de tipo atribuída àquela variável, enquanto que as funções que não retornam são representadas a partir do construtor de negação *TNeg [CPSMonoType]* sendo os argumentos dela definidos recursivamente. O contexto por sua vez, é um tipo que utiliza o *Data.Map* disponível no pacote *containers* para mapear uma variável para um tipo polimórfico. As substituições são representadas utilizando o mesmo *Data.Map*, onde desta vez é mapeado uma variável de tipo para um tipo monomórfico.

3.2.2 Traduções

Para facilitar os testes efetuados e ainda tornar o fluxo de execução mais direto, foram implementadas as funções de tradução de expressões e de tipos de acordo com as definições das Seções 2.5.1 e ??.

```
1 initialCont :: Id
2 initialCont = "k"
```

Código 3.2 – Continuação inicial

Para todas as computações, um contexto inicial precisa conter a continuação inicial. Este será o objeto a ter seu tipo inferido. Afim de praticidade, esta continuação será sempre a mesma, dada por ‘*k*’, como mostrado no Código 3.2.

```
1 callByName :: Expr -> Id -> FreshM Command
2 callByName (Var x) k = return $ Jump x [k]
3
4 callByName (Lam x e) k = do
5   k' <- freshCont
6   v <- freshVar
7   let jump = Jump k [v]
8   eCall <- callByName e k'
9   return $ Bind jump v [x, k'] eCall
```

```

10
11 callByName (App f e) k = do
12   fCont <- freshCont
13   fVar <- freshVar
14   argCont <- freshCont
15   argVar <- freshVar
16   fCall <- callByName f fCont
17   let appCmd = Jump fVar [argVar, k]
18   eCall <- callByName e argCont
19   let innerBind = Bind appCmd argVar [argCont] eCall
20   return $ Bind fCall fCont [fVar] innerBind
21
22 callByName (Let x b a) k = do
23   k' <- freshCont
24   bCall <- callByName b k'
25   aCall <- callByName a k
26   return $ Bind aCall x [k'] bCall
27
28
29 cbnExprTrans :: Expr -> Command
30 cbnExprTrans expr = evalState (callByName expr initialCont) (0, 0)

```

Código 3.3 – Tradução das expressões para CBN

No Código 3.3, são apresentadas as funções responsáveis para traduzir as expressões para CBN. O ponto de partida desta computação será a função `cbnExprTrans :: Expr → Command`, onde ela irá receber a expressão em cálculo- λ e retornará o comando traduzido. Sua responsabilidade é criar as mônadas de estado que farão o controle do índice das variáveis frescas necessárias e chamar as funções de tradução passando a continuação inicial.

Tomando como exemplo a função identidade em cálculo- λ ($\lambda x. x$), é possível perceber como os termos crescem em CPS. Isto torna o desenvolvimento diretamente neste cálculo não apropriado, mas ainda, nota-se que a implementação da função de tradução é bastante direta em relação a sua definição formal. Os únicos pontos de atenção são em relação à geração das variáveis frescas, mas que como foi dito anteriormente, não eram completamente necessários, visto que eles são ligados imediatamente. Ao traduzir então a função, tem-se que o equivalente em CPS é o mostrado no Código 3.4.

```

1 let v0(x, k0) =
2   x(k0)
3 in
4   k(v0)

```

Código 3.4 – Tradução da função identidade em CBN

Este comportamento fica ainda mais visível quando uma função um pouco maior é traduzida, por exemplo o numeral de Church dois ($\lambda f. \lambda x. f (f x)$). Sua tradução portanto é mostrada no Código 3.5.

```

1 let v0(f, k0) =
2   let v1(x, k1) =
3     let k2(v2) =
4       let v3(k3) =
5         let k4(v4) =
6           let v5(k5) =
7             x(k5)
8           in
9             v4(v5, k3)
10          in
11            f(k4)
12          in
13            v2(v3, k1)
14        in
15          f(k2)
16      in
17        k0(v1)
18 in
19   k(v0)

```

Código 3.5 – Tradução do numeral de Church “2” em CBN

De maneira semelhante, foram feitas as mesmas funções utilizadas no *call-by-name*, porém adaptadas para o CBV, respeitando as diferenças presentes na definição formal da função.

```

1 callByValue :: Expr -> Id -> FreshM Command
2 callByValue (Var x) k = return $ Jump k [x]
3
4 callByValue (Lam x e) k = do
5   k' <- freshCont
6   v <- freshVar
7   let bindBody = Jump k [v]
8   body <- callByValue e k'
9   return $ Bind bindBody v [x, k'] body
10
11 callByValue (App f e) k = do
12   fCont <- freshCont
13   fVar <- freshVar
14   argCont <- freshCont
15   argVar <- freshVar

```

```

16  fCall <- callByValue f fCont
17  eCall <- callByValue e argCont
18  let appCmd = Jump fVar [argVar, k]
19  let argBind = Bind eCall argCont [argVar] appCmd
20  return $ Bind fCall fCont [fVar] argBind
21
22 callByValue (Let x b a) k = do
23   k' <- freshCont
24   bCall <- callByValue b k'
25   aCall <- callByValue a k
26   return $ Bind bCall k' [x] aCall
27
28
29 cbvExprTrans :: Expr -> Command
30 cbvExprTrans expr = evalState (callByValue expr initialCont) (0, 0)

```

Código 3.6 – Tradução das expressões para CBV

A partir dessas diferenças nas definições, nota-se também particularidades nas traduções destas, por exemplo ao traduzir a mesma função identidade, em CBV, é obtido o resultado apresentado no Código 3.7.

```

1  let v0(x, k0) =
2      k0(x)
3  in
4      k(v0)

```

Código 3.7 – Tradução da função identidade em CBV

Neste exemplo da função identidade, pouca diferença entre as duas estratégias de avaliação pode ser notada. Isso se deve não ao tamanho da expressão, e sim dos elementos desta. Neste caso, há somente uma abstração lambda com uma variável. A seguir, é exibido novamente o numeral de Church dois, porém para o CBV, onde mais diferenças podem ser observadas. O motivo disto é os elementos da função, que diferentemente da identidade, conta com mais construtores para representá-la, exibido no Código 3.8.

```

1  let v0(f, k0) =
2      let v1(x, k1) =
3          let k2(v2) =
4              let k3(v3) =
5                  v2(v3, k1)
6              in
7                  let k4(v4) =
8                      let k5(v5) =
9                          v4(v5, k3)

```

```

10             in
11             k5(x)
12         in
13         k4(f)
14     in
15     k2(f)
16 in
17     k0(v1)
18 in
19     k(v0)

```

Código 3.8 – Tradução do numeral de Church “2” em CBV

Mostrado anteriormente no Código 3.4, a expressão CPS resultante da tradução da função identidade em CBN difere da mesma traduzida em CBV, presente no Código 3.7. O mesmo pode ser observado ao traduzir a função identidade tipada. Por exemplo, ao executar a função para a identidade em CBN, o tipo obtido é $\neg\neg(\neg\neg\alpha, \neg\alpha)$. Já em CBV, para a mesma função, tem-se que o tipo traduzido é $(\alpha, \neg\alpha)$.

Para raciocinar sobre os tipos do sistema, tem que ser levado em consideração o que estes representam, contradições. Ao tomar como exemplo o tipo resultante da função identidade em CPS a partir da tradução por valor (CBV), isto é, $(\alpha, \neg\alpha)^2$, deve-se pensar que este representa o absurdo de ter α como argumento e $\neg\alpha$ como continuação, ao mesmo tempo.

```

1 cbnTypeTranslation :: LambdaMonoType -> CPSPolyType
2 cbnTypeTranslation st =
3   let (transMono, freeVars) = runState (cbnTrans st) S.empty
4   in Forall (S.toList freeVars) transMono
5   where
6     cbnTrans :: LambdaMonoType -> State (S.Set String) CPSMonoType
7     cbnTrans = \case
8       Lambda.Typing.TVar varId -> do
9         modify (S.insert varId)
10        return $ TNeg [CPS.Typing.TVar varId]
11     TArr a b -> do
12       a' <- cbnTrans a
13       b' <- cbnTrans b
14       return $ TNeg [TNeg [TNeg [a'], b']]

```

Código 3.9 – Tradução dos tipos para CBN

Aqui no Código 3.9, a função responsável por traduzir um tipo utilizando a estratégia *call-by-name*, a função `cbnTypeTranslation :: LambdaMonoType → CPSPolyType` irá receber um

² Apesar de estar sendo traduzido o tipo da função identidade, como esta tradução é feita para os tipos simples, note que aqui, está sendo representado monomorficamente pois não há a generalização do α .

tipo simples no cálculo- λ simplesmente tipado, e retornar um tipo polimórfico em CPS. Perceba que na definição, o retorno era um tipo simples em CPS. Essa diferença é justificada ao se observar o corpo desta função, onde há a definição da função `cbvTrans`. Esta é quem efetivamente faz a computação, pois é nela que acontece o casamento de padrões para determinar o tipo sendo traduzido. Ainda, esta função retorna um tipo polimórfico pelo fato de que em momento posterior a essa tradução, a subtipagem do tipo traduzido e do tipo inferido precisa ser verificada.

```

1 cbvTypeTranslation :: LambdaMonoType -> CPSPolyType
2 cbvTypeTranslation st =
3   let (transMono, freeVars) = runState (cbvTrans st) S.empty
4       finalMono = TNeg [transMono]
5   in Forall (S.toList freeVars) finalMono
6   where
7     cbvTrans :: LambdaMonoType -> State (S.Set String) CPSMonoType
8     cbvTrans = \case
9       Lambda.Typing.TVar varId -> do
10        modify (S.insert varId)
11        return $ CPS.Typing.TVar varId
12     TArr a b -> do
13       a' <- cbvTrans a
14       b' <- cbvTrans b
15       return $ TNeg [a', TNeg [b']]

```

Código 3.10 – Tradução dos tipos para CBV

Para que essas duas funções de tradução de tipos (CBN e CBV) sejam usadas do mesmo modo, elas seguem a mesma assinatura. Seus comportamentos são o mesmo, a diferir somente nas diferenças das definições da função, isto é, como é feita a tradução.

O sistema de tipos proposto neste trabalho, entretanto, é polimórfico, ou seja, aqui há a adição de variáveis de tipos quantificadas. Sendo assim, esta função de tradução, se aplicando neste caso de uso, não necessariamente retornará o tipo mais geral de uma expressão, mas sempre um subtipo deste.

3.2.3 Inferência

A inferência de tipos é a etapa onde, em uma linguagem onde tipos são presentes, um termo tem seu tipado inferido sem anotação prévia. Isto é, sem explicitar o tipo de um termo, este tem seu tipo deduzido. Em um ambiente tipado com polimorfismo, a maior utilidade do inferidor de tipos é que este seja sempre o mais geral possível, tal que possa ser especializado para cada uso. O algoritmo de inferência proposto para este sistema de tipos é capaz de inferir o tipo mais geral possível, conforme explicado mais detalhadamente na Seção 3.1. Seu desenvolvimento foi direto como pode ser visto abaixo, no Código 3.11.

```

1 inferAtom :: Context -> Id -> TI CPSMonoType
2 inferAtom ctx x =
3   case readMaybe x :: Maybe Integer of
4     Just _ -> return TInt
5     Nothing -> case Map.lookup x ctx of
6       Just poly -> do
7         t <- instantiate poly
8         return t
9       Nothing -> throwError $ UnboundVariable x ctx
10
11
12 inferCommand :: Context -> Command -> TI Substitution
13 inferCommand ctx (Jump k xs) = do
14   t1 <- inferAtom ctx k
15   t2 <- mapM (inferAtom ctx) xs
16   s <- mgu t1 (TNeg t2)
17   return s
18
19 inferCommand ctx (Bind b y ys c) = do
20   paramTypes <- mapM (const freshTVar) ys
21   let ctxParams = extendContextWithParams ctx ys paramTypes
22   s1 <- inferCommand ctxParams c
23   let contType = applySubst s1 (TNeg paramTypes)
24   let ctxSubst = applySubstToContext s1 ctx
25   let sigma = generalize ctxSubst contType
26   let ctx' = Map.insert y sigma ctxSubst
27   s2 <- inferCommand ctx' b
28   return (composeSubst s2 s1)
29
30
31 inferWithCtx :: Command -> TI CPSPolyType
32 inferWithCtx cmd = do
33   initialType <- freshTVar
34   let ctx = Map.singleton initialCont (Forall [] initialType)
35
36   subst <- inferCommand ctx cmd
37   let ctx' = applySubstToContext subst ctx
38
39   case Map.lookup initialCont ctx' of
40     Just (Forall _ monoType) ->
41       let gen = generalize (Map.delete initialCont ctx') monoType

```

```

42     normalized = normalizePolyType gen
43   in return normalized
44   Nothing -> throwError (UnboundVariable initialCont ctx)

```

Código 3.11 – Função principal de Inferência

Muitas das funções necessárias para a inferência do CPS são iguais as do Damas-Milner. Por não ser o foco deste trabalho, explicações sobre estas serão omitidos. Desta forma, o processamento para se inferir o tipo das continuacões foi dividido em três principais funções.

A função `inferAtom :: Context → Id → TI CPSMonoType`, como seu nome e assinatura indica, é a função responsável por inferir os átomos do termo, isto é, buscar as variáveis do contexto e retorná-las caso sejam literais, instanciá-las se forem variáveis polimórficas, ou então retornar erro caso esta não esteja presente. Ao executar a função passando o contexto $\{x: \forall \alpha. \alpha\}$ e a variável x , a função irá instanciar uma nova variável de tipo β (se esta for a próxima ainda não utilizada) e retornar este tipo monomórfico. Ou ainda, caso a função seja chamada com o contexto $\{x: \forall \alpha. \alpha\}$ e a variável k , um erro `UnboundVariable` “ k ” $\{x: \forall \alpha. \alpha\}$ será exibido.

Já a função `inferCommand :: Context → Command → TI Substitution` é quem, a partir do contexto e comando, irá retornar a substituição que indica o tipo da continuação. Tal qual a função anterior, nenhuma dificuldade grande foi encontrada aqui. A inferência do salto foi a mais simples delas, onde cada linha da função se refere às premissas na definição, retornando então a unificação mais geral entre τ_1 e $\vec{\tau}_2$. Já para o *binding*, algumas funções auxiliares foram necessárias principalmente para manipular o contexto ao estender este em diferentes momentos. Além destas, funções para a generalização e por fim para a composição das substituições também foram precisas.

O ponto de partida da inferência, `inferWithCtx :: Command → TI CPSPolyType`, além de definir o contexto inicial com a continuação inicial k , recebendo um tipo α qualquer, esta função aplicará a substituição obtida na inferência e a aplicará no contexto, de modo que o contexto final esteja atualizado com o tipo inferido da continuação k . Após isto, é feita a normalização do tipo polimórfico da substituição, isto é, limpar as variáveis de tipo utilizadas durante o processo que não são mais necessárias, afim de promover consistência e facilitar o entendimento. Por exemplo, no caso onde o tipo polimórfico final seja $\forall \delta. \delta$, o retornado seria $\forall \alpha. \alpha$.

Uma etapa que não está diretamente relacionada com a inferência, e sim com os tipos em si, é a verificação da subtipagem do tipo traduzido e do tipo inferido. Isto é, uma verificação se o tipo traduzido é um subtipo do tipo inferido, indicando diretamente se o algoritmo de inferência foi implementado corretamente.

```

1 isSubtypeOfPoly :: CPSPolyType -> CPSPolyType ->
2   Either TypeError (Maybe Substitution)
3 isSubtypeOfPoly (Forall vars1 t1) (Forall vars2 t2) =
4   runTI $ do

```



```

5   fv1 <- replicateM (length vars1) freshTVar
6   let s1 = Map.fromList (zip vars1 fv1)
7   fv2 <- replicateM (length vars2) freshTVar
8   let s2 = Map.fromList (zip vars2 fv2)
9   pure $ isSubtypeOf (applySubst s1 t1) (applySubst s2 t2)
10
11
12 isSubtypeOf :: CPSMonoType -> CPSMonoType -> Maybe Substitution
13 isSubtypeOf t1 t2 = match t1 t2 Map.empty
14   where
15     match :: CPSMonoType -> CPSMonoType -> Substitution ->
16       Maybe Substitution
17     match (TVar a) t subst =
18       case Map.lookup a subst of
19         Just tExisting ->
20           if tExisting == t
21             then Just subst
22             else Nothing
23         Nothing ->
24           if occursCheck a t
25             then Nothing
26             else Just (Map.insert a t subst)
27     match TInt TInt subst = Just subst
28     match (TNeg ts1) (TNeg ts2) subst
29       | length ts1 == length ts2 =
30         foldM
31           (\s (t1', t2') -> match t1' t2' s)
32           subst
33           (zip ts1 ts2)
34       | otherwise = Nothing
35     match _ _ _ = Nothing

```

Código 3.12 – Verificação de Subtipagem

O algoritmo de verificação de subtipos do Código 3.12 procura uma substituição S tal que, ao aplicá-la em um dos tipos, ele se torne o outro. Por exemplo, ao analisar os tipos $\tau_1 = \alpha \rightarrow \alpha$ e $\tau_2 = \neg\alpha \rightarrow \neg\alpha$, é possível encontrar uma substituição $S = \{ \alpha \mapsto \neg\alpha \}$, tal que $S\tau_1$ resulte em τ_2 . A partir disto, pode ser dito que o tipo τ_2 é um subtipo de τ_1 , onde o algoritmo retornaria com sucesso a substituição S . Já olhando para outro exemplo, onde $\tau_1 = \alpha \rightarrow \neg\alpha$ e $\tau_2 = \alpha \rightarrow \alpha$, não é possível encontrar uma substituição S tal que $S\tau_1$ seja τ_2 . Desta forma, o algoritmo não retornaria uma substituição, falhando assim a verificação de subtipos.

3.2.4 Geração de Código

Uma vez que as provas de completude e consistência não estavam no escopo deste trabalho e a validação do algoritmo se deu por meio de testes, um conjunto considerável de testes foi construído. Para isolar os testes, ou seja, testar as funções de maneira independente para garantir que, se a inferência apresentasse algum erro, fosse certo que o erro estaria na inferência e não por conta de uma tradução incorreta, dois teoremas apresentados em (PLOTKIN, 1975) foram utilizados.

Este teorema, chamado de Teorema da Simulação, válido tanto para *call-by-name* quanto para *call-by-value*, afirma que, dado um programa M em cálculo- λ , o resultado de sua computação há de ser o mesmo que o resultado da computação da tradução para CPS com a função identidade. Ou seja, simulando a execução do programa feito em cálculo- λ no cálculo de continuuações. Ou então, formalmente, $Eval(M) = Eval([M](\lambda x. x))$, onde a função $Eval$ é responsável por avaliar uma expressão, efetivamente a computando.

Foi implementada a simulação do cálculo- λ no cálculo de continuuações a partir das funções de tradução para programas que computam os numerais de Church. Isto é feito computando a função λ onde, é incrementado o valor que representa o número de aplicações feitas (essencialmente como um numeral de Church é computado), passando como continuação a função identidade. Fazendo com que assim, a função lambda que representa um numeral de church é simulada pelo cálculo de continuuações.

```

1  -- Auto-generated by Main.hs
2  -- For input {{ Input file name }}
3  -- This code works only for Church Encoding expressions
4  -- It will compute the ...
5  -- ... generated CPS translated expression Church encoding
6
7  -----
8
9  cbn k = -- {{ CBN Translated Expression }}
10
11  cbv k = -- {{ CBV Translated Expression }}
12
13  -----
14
15  inc_cbv :: (Int, Int -> r) -> r
16  inc_cbv (n, k) = k (1 + n)
17
18  test_cbv =
19      cbv (\f -> f (inc_cbv, \x -> x (0, id)))
20
21  thunk :: a -> (a -> r) -> r

```

```

22 thunk x k = k x
23
24 inc_cbn :: (Int -> r, Int) -> r
25 inc_cbn (k, n) =
26     inc_cbv (n, k)
27
28 test_cbn =
29     cbn (\f -> f (thunk inc_cbn, \x -> x (id, 0)))
30
31 main :: IO ()
32 main = do
33     print $ (test_cbn, test_cbv)

```

Código 3.13 – Geração de código para computação de numerais de Church

O código Haskell gerado, apresentado no Código 3.13 pode ser dividido em três partes principais, o cabeçalho de caráter informativo, que explicita o programa em cálculo- λ de entrada para ter gerado aquele programa em CPS. Em seguida, há a definição das funções *cbn* e *cbv*, ou seja, os programas correspondente em CPS para as duas traduções daquela entrada. Por fim, a última parte é responsável pela computação do numeral de Church, as funções definidas irão calcular o número representado pelas expressões em CPS e retornar por fim uma tupla contendo o resultado do calculado pelo *call-by-name* e *call-by-value* respectivamente.

Ao se traduzir o numeral de Church 0 representado em cálculo- λ por $\lambda f.\lambda x. x$, para CBN e CBV, respectivamente nos Códigos 3.14 e 3.15, tem-se:

```

1 let v0(f, k0) =
2   let v1(x, k1) =
3     x(k1)
4   in
5     k0(v1)
6 in
7   k(v0)

```

Código 3.14 – Tradução do numeral de Church “0” em CBN

```

1 let v0(f, k0) =
2   let v1(x, k1) =
3     k1(x)
4   in
5     k0(v1)
6 in
7   k(v0)

```

Código 3.15 – Tradução do numeral de Church “0” em CBV

Desta forma, o código gerado ao se traduzir esta função, é ilustrado Código 3.16.

```

1  -- Auto-generated by Main.hs
2  -- For input church-zero
3  -- This code works only for Church Encoding expressions
4  -- It will compute the ...
5  -- ... generated CPS translated expression Church encoding
6
7  -----
8
9  cbn k = let v0(f, k0) = let v1(x, k1) = x(k1) in k0(v1) in k(v0)
10
11  cbv k = let v0(f, k0) = let v1(x, k1) = k1(x) in k0(v1) in k(v0)
12
13  -----
14
15  inc_cbv :: (Int, Int -> r) -> r
16  inc_cbv (n, k) = k (1 + n)
17
18  test_cbv =
19      cbv (\f -> f (inc_cbv, \x -> x (0, id)))
20
21  thunk :: a -> (a -> r) -> r
22  thunk x k = k x
23
24  inc_cbn :: (Int -> r, Int) -> r
25  inc_cbn (k, n) =
26      inc_cbv (n, k)
27
28  test_cbn =
29      cbn (\f -> f (thunk inc_cbn, \x -> x (id, 0)))
30
31  main :: IO ()
32  main = do
33      print $ (test_cbn, test_cbv)

```

Código 3.16 – Código gerado ao traduzir o numeral de Church “0”

Ao executar o código e chamar a função main deste programa, o resultado obtido é justamente a computação do numeral para as duas traduções, ou seja, (0, 0).

3.2.5 Fluxo Principal

O fluxo completo de execução do programa principal contempla todas as funções apresentadas nesta seção, com a adição de funções auxiliares. Essas são aplicadas em sequência, de modo a realizar uma série de ações.

Inicialmente, é passado o caminho de um arquivo contendo um programa em cálculo- λ com adição do 'let'. O conteúdo então é processado pelo *parser* e representado pelos tipos de dados algébricos para o cálculo lambda. Uma vez que o programa já está sendo representado pelos ADTs, e ainda tem seu tipo inferido, é possível iniciar o processamento descrito pelas funções apresentadas. A primeira delas é a tradução para CPS tanto em *call-by-name* quanto em *call-by-value*, as exibindo logo em seguida. Com as traduções da expressão feitas, o código Haskell já pode ser gerado, salvando assim no diretório output com mesmo nome do arquivo de entrada. Como passo posterior, tem-se a tradução dos tipos para ambas as estratégias de avaliação. Os passos finais envolvem a inferência de ambas as traduções, juntamente da verificação de subtipagem, onde esta informará se o tipo traduzido é um subtipo do inferido.

```
$ cabal run
Input file path:
> input/church-zero.in
Expression:
 $\lambda f. \lambda x. x$ 
Type:
 $\alpha \rightarrow \beta \rightarrow \beta$ 
Call-by-Name Translation:
Command:
let v0(f, k0) =
  let v1(x, k1) =
    x(k1)
  in
    k0(v1)
in
  k(v0)
Expected Continuation Type:
 $\forall \alpha, \beta. \neg(\neg \alpha, \neg(\neg \beta, \neg \beta))$ 
Inferred Continuation Type:
 $\forall \alpha, \beta. \neg(\alpha, \neg(\neg \beta, \beta))$ 
Do the types match?
Yes

Call-by-Value Translation:
Command:
let v0(f, k0) =
```

```

    let v1(x, k1) =
        k1(x)
    in
        k0(v1)
in
    k(v0)
Expected Continuation Type:
 $\forall \alpha, \beta. \neg\neg(\alpha, \neg\neg(\beta, \neg\beta))$ 
Inferred Continuation Type:
 $\forall \alpha, \beta. \neg\neg(\alpha, \neg\neg(\beta, \neg\beta))$ 
Do the types match?
Yes

CPS expression saved in output/church-zero.hs

```

Código 3.17 – Execução do programa principal

Ao executar o programa do Código 3.17 com o comando `cabal run`, passando também o arquivo de entrada `input/church-zero.in`, é processado e exibida todas as informações que foram citadas anteriormente, inclusive a geração do código Haskell em `output/church-zero.hs`. É possível perceber que na saída do programa, é mostrado o tipo traduzido (na sequência da mensagem “*Expected Continuation Type:*”) e o tipo inferido (que sucede a mensagem “*Inferred Continuation Type:*”). Logo em seguida, o questionamento “*Do the types match?*” é o trecho da saída que compete à subtipagem, retornando “*Yes*” caso esse seja um subtipo deste, o que indica uma inferência compatível com a tradução ou “*No*” caso a verificação falhe, indicando uma inferência incorreta.

```

$ ghc output/church-zero.hs
[1 of 2] Compiling Main                ( output/church-zero.hs, ... )
[2 of 2] Linking output/church-zero
$ ./output/church-zero
(0,0)

```

Código 3.18 – Execução do programa gerado

Ainda, a compilação e execução do código Haskell gerado pode ser conferida no Código 3.18 acima, ilustrando exatamente o comportamento detalhado anteriormente.

4 RESULTADOS

Uma vez que todo o fluxo do programa foi exibido, seus resultados podem ser apresentados e compreendidos. Diversos testes foram executados onde, a partir da função de tradução de tipos, puderam ter seus tipos verificados para validar a implementação. Para tal, estão disponibilizados no repositório do projeto, arquivos de entrada com funções contendo diferentes características. Afim de observar o comportamento do programa abrangendo uma maior gama de opções, algumas das funções testadas são extensas e contam com combinações de regras do cálculo- λ simplesmente tipado.

4.1 COMBINADOR S

Um desses é o combinador S ($\lambda x. \lambda y. \lambda z. x z (y z)$), que apesar de não ser um termo extenso, este utiliza de uma combinação dos três construtores (variáveis, abstrações e aplicações) para que o termo seja obtido. O tipo deste, é representado por $(\alpha \rightarrow \beta \rightarrow \gamma) \rightarrow (\alpha \rightarrow \beta) \rightarrow \alpha \rightarrow \gamma$ e suas traduções, tanto em CBN quanto em CBV são grandes demais para serem aqui colocadas, mas estas podem ser encontradas no código fonte do repositório mencionado. A tradução do seu tipo entretanto, para CBN, é apresentada a seguir.

$$\forall \alpha, \beta, \gamma. \neg \neg (\neg \neg \neg (\neg \neg \alpha, \neg \neg (\neg \neg \beta, \neg \gamma)), \neg \neg (\neg \neg \neg (\neg \neg \alpha, \neg \beta), \neg \neg (\neg \neg \alpha, \neg \gamma)))$$

Enquanto que, a partir do tipo inferido a seguir para a mesma estratégia de avaliação, uma substituição S tal que ao aplicá-la no tipo traduzido torne-se o inferido, é $S = \{ \alpha \mapsto \neg \alpha, \beta \mapsto \neg \beta, \gamma \mapsto \neg \gamma \}$, validando assim a inferência para este termo.

$$\forall \alpha, \beta, \gamma. \neg \neg (\neg \neg \neg (\neg \alpha, \neg \neg (\neg \beta, \gamma)), \neg \neg (\neg \neg \neg (\neg \alpha, \beta), \neg \neg (\neg \alpha, \gamma)))$$

Um comportamento semelhante pode ser percebido para a tradução por *call-by-value*, onde respectivamente é apresentado a seguir a tradução e o resultado da inferência.

$$\forall \alpha, \beta, \gamma. \neg \neg (\neg (\alpha, \neg \neg (\beta, \neg \gamma)), \neg \neg (\neg (\alpha, \neg \beta), \neg \neg (\alpha, \neg \gamma)))$$

Neste caso, a substituição S que satisfaz a condição de subtipagem é tal que $S = \{ \gamma \mapsto \neg \gamma \}$, tornando válida assim a inferência para este termo.

$$\forall \alpha, \beta, \gamma. \neg \neg (\neg (\alpha, \neg \neg (\beta, \gamma)), \neg \neg (\neg (\alpha, \neg \beta), \neg \neg (\alpha, \gamma)))$$

4.2 SOMA

O próximo exemplo apresentado é a função de soma de 2 e 3 feita com os numerais de Church $((\lambda n. \lambda m. \lambda f. \lambda x. n f (m f x)) (\lambda a. \lambda b. a (a b)) (\lambda c. \lambda d. c (c (c d))))$. Seu tipo, é representado por $((\alpha \rightarrow \alpha) \rightarrow \alpha \rightarrow \alpha)$. A tradução do tipo para CBN é dado por:

$$\forall \alpha. \neg \neg (\neg \neg \neg (\neg \neg \alpha, \neg \alpha), \neg \neg (\neg \neg \alpha, \neg \alpha))$$

E o inferido também para CBN, onde a substituição S que satisfaz a subtipagem é tal que $S = \{ \alpha \mapsto \neg\alpha \}$:

$$\forall\alpha. \neg\neg(\neg\neg\neg(\neg\alpha, \alpha), \neg\neg(\neg\alpha, \alpha))$$

Enquanto que para CBV, o tipo traduzido é:

$$\forall\alpha. \neg\neg(\neg(\alpha, \neg\alpha), \neg\neg(\alpha, \neg\alpha))$$

O inferido portanto, sendo que a substituição S que satisfaz a subtipagem neste caso é a substituição trivial $S = \{ \alpha \mapsto \alpha \}$:

$$\forall\alpha. \neg\neg(\neg(\alpha, \neg\alpha), \neg\neg(\alpha, \neg\alpha))$$

Por fim, ao executar o código gerado, obtém-se o resultado (5, 5), indicando uma correta tradução da expressão de entrada.

4.3 MULTIPLICAÇÃO

Outro exemplo é o da multiplicação de 6 e 8, que o termo lambda por si é grande demais para ser exibido aqui. A função lambda responsável pela multiplicação de dois argumentos, $(\lambda m. \lambda n. \lambda f. \lambda x. m (n f) x)$ é extenso o bastante para seu sucesso dar uma noção boa de que o código está correto e que o algoritmo é capaz de inferir o tipo corretamente das mais diversas expressões. Assim como o exemplo da função de soma, o tipo desta expressão é: $((\alpha \rightarrow \alpha) \rightarrow \alpha \rightarrow \alpha)$. A tradução do tipo para CBN é dado por:

$$\forall\alpha. \neg\neg(\neg\neg\neg(\neg\neg\alpha, \neg\alpha), \neg\neg(\neg\neg\alpha, \neg\alpha))$$

E o inferido também para CBN, onde a substituição S que satisfaz a subtipagem é tal que $S = \{ \alpha \mapsto \neg\alpha \}$:

$$\forall\alpha. \neg\neg(\neg\neg\neg(\neg\alpha, \alpha), \neg\neg(\neg\alpha, \alpha))$$

Enquanto que para CBV, o tipo traduzido é:

$$\forall\alpha. \neg\neg(\neg(\alpha, \neg\alpha), \neg\neg(\alpha, \neg\alpha))$$

O inferido portanto, sendo que a substituição S que satisfaz a subtipagem neste caso é a substituição trivial $S = \{ \alpha \mapsto \alpha \}$:

$$\forall\alpha. \neg\neg(\neg(\alpha, \neg\alpha), \neg\neg(\alpha, \neg\alpha))$$

É possível perceber que, os resultados obtidos são os mesmos da função de soma. Isso se deve ao fato de que, por mais que a expressão seja diferente, as duas possuem o mesmo tipo. Então, o algoritmo infere que as duas possuem o mesmo tipo.

Por fim, ao executar o código gerado, obtém-se o resultado (48, 48), indicando uma correta tradução da expressão de entrada.

4.4 IDENTIDADE COM ‘LET’

Ainda que a tradução do ‘let’ tenha sido apresentada e discutida, esta possui um problema relacionada à tipagem quanto ao *call-by-value*. Para melhor compreender esta questão, é necessário entender fundamentalmente a diferença entre as estratégias de avaliação por nome e por valor. A por nome, também conhecida como avaliação preguiçosa, só irá avaliar a expressão no momento em que esta for necessária, desta forma, caso tenha alguma expressão que não seja utilizada, esta nem mesmo será computada. Enquanto que a por valor não, ao invés disso, ela avalia toda expressão no início, sem se importar se será utilizada ou não, desta forma, ainda que uma expressão não seja utilizada, ela será calculada.

A função em questão é a identidade utilizando o ‘let’, ($\text{let id} = \lambda x. x \text{ in id id}$). Seu tipo é $\alpha \rightarrow \alpha$, enquanto que sua tradução em CBN é dada pelo Código 4.1.

```

1 let id(k0) =
2   let v0(x, k1) =
3     x(k1)
4   in
5     k0(v0)
6 in
7   let k2(v1) =
8     let v2(k3) =
9       id(k3)
10    in
11      v1(v2, k)
12  in
13    id(k2)

```

Código 4.1 – Tradução em CBN da identidade com ‘let’

A tradução do tipo para CBN é dada por:

$$\forall \alpha. \neg \neg (\neg \neg \alpha, \neg \alpha)$$

E o inferido também para CBN, onde a substituição S que satisfaz a subtipagem é tal que $S = \{ \alpha \mapsto \neg \alpha \}$:

$$\forall \alpha. \neg \neg (\neg \alpha, \alpha)$$

Já ao olhar para o resultado em CBV, temos um problema de tipagem. A tradução da expressão, que num primeiro momento não há nenhum problema aparente, é dada pelo Código 4.2.

```

1 let k0(id) =
2   let k2(v1) =
3     let k3(v2) =
4       v1(v2, k)

```

```

5      in
6      k3(id)
7      in
8      k2(id)
9  in
10     let v0(x, k1) =
11         k1(x)
12     in
13         k0(v0)

```

Código 4.2 – Tradução em CBV da identidade com ‘let’

Ao investigar mais a fundo porém, é possível ser notado na linha 1 que, como o `id` é passado por parâmetro para a continuação `k`, durante o processo de inferência, esta será inserida no contexto como sendo um tipo polimórfico. Em momento posterior entretanto, na inferência da expressão `id id`, a expressão `id` é assumida como sendo um tipo monomórfico. É neste instante então que é feita a verificação do *occurs check* para garantir que tipos cíclicos (ou seja, um tipo estar contido em outro, tornando assim impossível a unificação destes) não sejam permitidos, onde esta falha, retornando assim o erro. Desta forma, o algoritmo falha, identificando um erro de *OccursCheck*, que será apresentado a seguir:

$$\text{OccursCheck: } \beta \text{ in } \neg(\beta, \alpha)$$

Este erro indica que, ao tentar encontrar uma unificação para as variáveis de tipo, foi encontrado que uma delas já estava presente na outra, neste caso, que o β já ocorria em $\neg(\beta, \alpha)$. Como isto torna impossível que uma substituição seja suficiente, de modo a gerar um loop infinito caso seja tentado, um erro é retornado.

4.5 AUTOAPLICAÇÃO MAL-TIPADA

Este exemplo demonstra um caso clássico de termo não tipável no cálculo lambda simplesmente tipado: a autoaplicação $(\lambda x. x x)$. No sistema de tipos original, a expressão requer que o tipo de x seja ao mesmo tempo uma função (para ser aplicada) e o argumento dessa função, levando a uma contradição. O erro de tipo obtido é:

$$\text{Cannot unify } \{ \alpha \} \text{ with } \{ \alpha \rightarrow \beta \}$$

A tradução para CPS em *call-by-name* é ilustrada pelo Código 4.3.

```

1  let v0(x, k0) =
2      let k1(v1) =
3          let v2(k2) =
4              x(k2)
5          in

```

```

6           v1(v2, k0)
7   in
8       x(k1)
9 in
10    k(v0)

```

Código 4.3 – Tradução em CBN da autoaplicação

Durante a inferência de tipos no sistema CPS, o algoritmo falha com:

$$\text{OccursCheck} : \varepsilon \text{ in } \neg\neg(\neg\varepsilon, \gamma)$$

Já na estratégia *call-by-value*, a tradução é apresentada no Código 4.4.

```

1 let v0(x, k0) =
2   let k1(v1) =
3     let k2(v2) =
4       v1(v2, k0)
5     in
6       k2(x)
7   in
8     k1(x)
9 in
10  k(v0)

```

Código 4.4 – Tradução em CBV da autoaplicação

E o erro de inferência correspondente:

$$\text{OccursCheck} : \beta \text{ in } \neg(\beta, \gamma)$$

Em ambos os casos, o sistema CPS apresenta erro na tipagem, refletindo a não tipagem do termo original.

4.6 COMBINADOR Y

O combinador Y ($\lambda f. (\lambda x. f(x x)) (\lambda x. f(x x))$) é outro termo não tipável no cálculo lambda simplesmente tipado, essencial para expressar recursão em cálculos não tipados. O erro de unificação é semelhante ao caso anterior:

$$\text{Cannot unify } \{ \beta \} \text{ with } \{ \beta \rightarrow \gamma \}$$

A tradução em *call-by-name* é:

```

1 let v0(f, k0) =
2   let k1(v1) =
3     let v2(k2) =

```

```

4      let v8(x, k8) =
5          let k9(v9) =
6              let v10(k10) =
7                  let k11(v11) =
8                      let v12(k12) =
9                          x(k12)
10                     in
11                         v11(v12, k10)
12                 in
13                     x(k11)
14             in
15                 v9(v10, k8)
16         in
17             f(k9)
18     in
19         k2(v8)
20 in
21     v1(v2, k0)
22 in
23     let v3(x, k3) =
24         let k4(v4) =
25             let v5(k5) =
26                 let k6(v6) =
27                     let v7(k7) =
28                         x(k7)
29                 in
30                     v6(v7, k5)
31             in
32                 x(k6)
33         in
34             v4(v5, k3)
35     in
36         f(k4)
37 in
38     k1(v3)
39 in
40     k(v0)

```

Código 4.5 – Tradução em CBN do combinador Y

A inferência de tipos no sistema CPS para CBN gera:

$$\text{OccursCheck} : \lambda \text{ in } \neg\neg(\neg\lambda, \iota)$$

Para *call-by-value*, a tradução é a seguinte:

```

1 let v0(f, k0) =
2   let k1(v1) =
3     let k2(v2) =
4       v1(v2, k0)
5     in
6       let v8(x, k8) =
7         let k9(v9) =
8           let k10(v10) =
9             v9(v10, k8)
10          in
11            let k11(v11) =
12              let k12(v12) =
13                v11(v12, k10)
14              in
15                k12(x)
16            in
17              k11(x)
18          in
19            k9(f)
20        in
21          k2(v8)
22    in
23      let v3(x, k3) =
24        let k4(v4) =
25          let k5(v5) =
26            v4(v5, k3)
27          in
28            let k6(v6) =
29              let k7(v7) =
30                v6(v7, k5)
31              in
32                k7(x)
33            in
34              k6(x)
35          in
36            k4(f)
37        in
38          k1(v3)
39    in
40      k(v0)

```

Código 4.6 – Tradução em CBV do combinador Y

Com o seguinte erro de inferência:

$$\text{OccursCheck} : \zeta \text{ in } \neg(\zeta, \neg t)$$

Estes resultados demonstram que o sistema de tipos proposto é consistente com o cálculo lambda tradicional, identificando corretamente termos não tipáveis através de erros de ocorrência durante a unificação.

5 CONCLUSÃO

Uma vez identificado ao menos um programa que não é capaz de inferir tipos para a tradução via ‘let’ da expressão, o seguinte teorema pode ser definido.

Teorema 1 (A tradução para CBV não preserva tipos). *A tradução do cálculo- λ para o cálculo de continuções em CBV não preserva tipos.*

Prova do Teorema 1. Podemos observar pelo contra exemplo: $\text{let id} = \lambda x. x \text{ in id id}$. Este termo não é tipável no cálculo de CPS polimórfico proposto. ■

Ainda, durante o desenvolvimento, ao se deparar com o erro de tipagem do ‘let’ no CBV, o pensamento inicial foi de que alguma confusão quanto a implementação da tradução havia sido cometida. Somente depois de muita análise e raciocínio a respeito do ocorrido foi compreendido que não se tratava de um erro de implementação, e sim de uma prova de que a tradução para CBV não preserva a tipagem dos termos.

Todos os testes realizados para CBN funcionaram, assim, é um forte indício que a tradução para este preserva a tipagem das expressões lambda. Desta forma, é definida a seguinte conjectura.

Conjectura 1 (A tradução para CBN preserva tipos). *A tradução do cálculo- λ para o cálculo de continuções em CBN preserva tipos.*

Por mais que a implementação em si não tenha sido complicada de modo geral, algumas dificuldades mais teóricas se mantiveram latentes durante boa parte do desenvolvimento. O mais incômodo deles, refere-se a falta de materiais de referência mais didáticos. A maioria deles, se não todos, apresenta as continuções utilizando conceitos sem previamente contextualizá-los. É esperado que o leitor tenha amplo conhecimento a respeito de Teoria de Tipos e Teoria das Categorias, o que dificulta o entendimento daqueles que não o tem.

Ao chegar no resultado de uma inferência, seja executando o algoritmo via código ou no papel, tem-se a prova do absurdo de uma continuação. Um dos pontos mais iniciais de dúvidas era referente ao tipo desta continuação. Seja a função identidade traduzida em CBV, ou seja, com tipo $\forall \alpha. \neg \neg(\neg \alpha, \alpha)$, a dificuldade que perdurava era como raciocinar sobre ele. O que contribui para essa barreira é o fato dos tipos deste sistema serem duais dos tipos no Sistema de Damas-Milner, precisando mudar a maneira de pensar para raciocinar sobre a tipagem.

5.1 CONSIDERAÇÕES FINAIS

As IRs são muito importantes na compilação, elas permitem otimizar muitos processos para assim tornar mais eficiente o código. Os sistemas de tipos permitem provas de propriedades de um sistema, garantindo que programas tenham comportamento esperado. Por exemplo, ao utilizar um sistema de tipos, é possível identificar um uso incorreto de uma função que opera

sobre valores numéricos quando é passado um argumento de tipo incompatível, impedindo assim sua execução para evitar comportamentos inesperados. Sendo assim, pode-se dizer que o sistema de tipos adiciona uma camada de segurança ao programa.

Uma das principais motivações para desenvolver um sistema de tipos para o CPS é adicionar uma camada de segurança na representação intermediária, estendendo a verificação de tipos para etapas posteriores do processo de compilação, como na fase de ligação de módulos (chamada de *linking*). Considere um cenário onde um programa possua milhares de arquivos fonte com IRs não tipadas, as propriedades de tipos das funções definidas em um arquivo só podem ser verificadas contra outros módulos ou arquivos se estes forem compilados conjuntamente — o que se torna inviável em programas de grande porte. Com uma IR tipada, se uma função f definida em um módulo espera receber um inteiro mas é invocada em outro módulo com um argumento do tipo *string*, o sistema de tipos identificará esse erro durante o (*linking*), sem necessidade de recompilar o módulo que contém f ou os módulos que a utilizam.

Mesmo que o sistema não funcione para o ‘let’ no CBV, a tradução está correta, onde isso significa respeitar o teorema da simulação. Isto é, no caso de uma implementação não tipada da tradução, não haveria problemas em simular a expressão lambda no cálculo de continuações. Como a tradução para CBV não preserva tipos, um dos trabalhos futuros incluiria a correção deste. Para tal, seria necessário propor um outro sistema de tipos polimórfico para o cálculo de continuações onde sejam distinguidos tipos e cotipos dentre os argumentos para que o algoritmo de inferência saiba quando utilizar um ou outro.

Outro trabalho futuro são as provas de consistência e completude do sistema de tipos proposto em relação ao algoritmo de inferência. Neste trabalho foram executados alguns casos de testes que demonstraram empiricamente que o algoritmo está correto. Em conjunto, um avanço seria a criação de um compilador que faça uso desta IR tipada, colocando em prática toda a teoria aqui apresentada.

REFERÊNCIAS

- AHO, Alfred V et al. **Compiladores: Princípios, técnicas e ferramentas**. 2th. ed. São Paulo, SP, Brasil: Pearson Education, 2008. Citado na página 17.
- APPEL, Andrew W. **Compiling with continuations**. USA: Cambridge University Press, 1992. ISBN 0521416957. Citado 3 vezes nas páginas 17, 20 e 31.
- BOWMAN, William J. et al. Type-preserving cps translation of σ and π types is not not possible. **Proceedings of the ACM on Programming Languages**, ACM, v. 2, n. POPL, p. 1–33, 2018. Disponível em: <<https://doi.org/10.1145/3158110>>. Citado na página 15.
- CHURCH, Alonzo. A set of postulates for the foundation of logic. **Annals of mathematics**, JSTOR, p. 346–366, 1932. Citado na página 18.
- CHURCH, Alonzo. A formulation of the simple theory of types. **J. Symb. Log.**, Cambridge University Press (CUP), v. 5, n. 2, p. 56–68, 1940. Citado na página 24.
- COOPER, Keith D; TORCZON, Linda. **Contruindo Compiladores**. 2th. ed. [S.l.]: Elsevier, 2014. Citado na página 17.
- COQUAND, Thierry. Type Theory. In: ZALTA, Edward N.; NODELMAN, Uri (Ed.). **The Stanford Encyclopedia of Philosophy**. [S.l.]: Metaphysics Research Lab, Stanford University, 2022. Citado na página 22.
- COQUAND, Thierry; HUET, Gérard. The calculus of constructions. **Information and Computation**, v. 76, n. 2, p. 95–120, 1988. Citado na página 25.
- DAMAS, Luis. **Type assignment in programming languages**. Tese (Doutorado) — University of Edinburgh, 1984. Citado na página 26.
- DAMAS, Luis; MILNER, Robin. **Principal type-schemes for functional programs**. Tese (Doutorado) — University of Edinburgh, Scotland, 1982. Citado 4 vezes nas páginas 25, 27, 28 e 29.
- FLANAGAN, Cormac et al. The essence of compiling with continuations. **SIGPLAN Not.**, Association for Computing Machinery, New York, NY, USA, v. 28, n. 6, p. 237–247, jun 1993. ISSN 0362-1340. Disponível em: <<https://doi.org/10.1145/173262.155113>>. Citado 2 vezes nas páginas 13 e 20.
- KENNEDY, Andrew. Compiling with continuations, continued. In: . [S.l.]: Association for Computing Machinery, 2007. Citado na página 13.
- MILNER, Robin. A theory of type polymorphism in programming. In: **Journal of Computer and System Sciences**. [S.l.]: Elsevier, 1978. v. 17, n. 3, p. 348–375. Citado 2 vezes nas páginas 25 e 26.
- MORRISETT, Greg et al. From system f to typed assembly language. **ACM Trans. Program. Lang. Syst.**, Association for Computing Machinery, New York, NY, USA, v. 21, n. 3, p. 527–568, maio 1999. ISSN 0164-0925. Disponível em: <<https://doi.org/10.1145/319301.319345>>. Citado 2 vezes nas páginas 13 e 15.
- MUCHNICK, Steven S. **Advanced Compiler Design and Implementation**. Oxford, England: Morgan Kaufmann, 1997. Citado na página 18.

NECULA, George C. Proof-carrying code. In: **Proceedings of the 24th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages**. New York, NY, USA: Association for Computing Machinery, 1997. (POPL '97), p. 106–119. ISBN 0897918533. Disponível em: <<https://doi.org/10.1145/263699.263712>>. Citado na página 14.

PIERCE, Benjamin C. **Types and Programming Languages**. [S.l.]: The MIT Press, 2002. Citado 3 vezes nas páginas 23, 24 e 25.

PLOTKIN, G.D. Call-by-name, call-by-value and the λ -calculus. **Theoretical Computer Science**, v. 1, n. 2, p. 125–159, 1975. Citado 3 vezes nas páginas 13, 33 e 49.

RIBEIRO, Rodrigo; CAMARÃO, Carlos. A mechanized textbook proof of a type unification algorithm. In: CORNÉLIO, Márcio; ROSCOE, Bill (Ed.). **Formal Methods: Foundations and Applications**. Cham: Springer International Publishing, 2016. p. 127–141. ISBN 978-3-319-29473-5. Citado 2 vezes nas páginas 29 e 30.

SHAO, Zhong; LEAGUE, Christopher; MONNIER, Stefan. Implementing typed intermediate languages. In: **Proceedings of the 1998 ACM SIGPLAN International Conference on Functional Programming (ICFP'98)**. Baltimore, Maryland, USA: ACM Press, 1998. Also available as Yale University Technical Report. Disponível em: <<http://flint.cs.yale.edu/flint/publications/icfp98.ps>>. Citado na página 15.

SILVA, Rafael Castro Gonçalves. **Uma Certificação em Coq do Algoritmo W Monádico**. Dissertação (Mestrado) — UDESC, 2019. Citado 2 vezes nas páginas 27 e 30.

SØNDERGAARD, Harald; SESTOFT, Peter. Referential transparency, definiteness and unfoldability. **Acta Inform.**, v. 27, n. 6, 1990. Citado na página 14.

THIELECKE, Hayo. **Categorical Structure of Continuation Passing Style**. 1997. Citado 5 vezes nas páginas 13, 14, 31, 32 e 33.

TORRENS, Paulo; ORCHARD, Dominic; VASCONCELLOS, Cristiano. On the operational theory of the cps-calculus: Towards a theoretical foundation for irs. **Proc. ACM Program. Lang.**, 2024. Citado 4 vezes nas páginas 13, 31, 32 e 34.

TORRENS, Paulo Henrique. **Um Cálculo de Continuações com Tipos Dependentes**. Dissertação (Mestrado) — UDESC, 2019. Citado na página 22.

VASCONCELLOS, Cristiano Damiani. **Inferência de Tipos com Suporte para Sobrecarga Baseada no Sistema CT**. Tese (Doutorado) — Universidade Federal de Minas Gerais, 2004. Citado na página 29.