

Compiladores

Técnicas e Ferramentas

O que são compiladores?

`printf("Hello");` →

Compilador

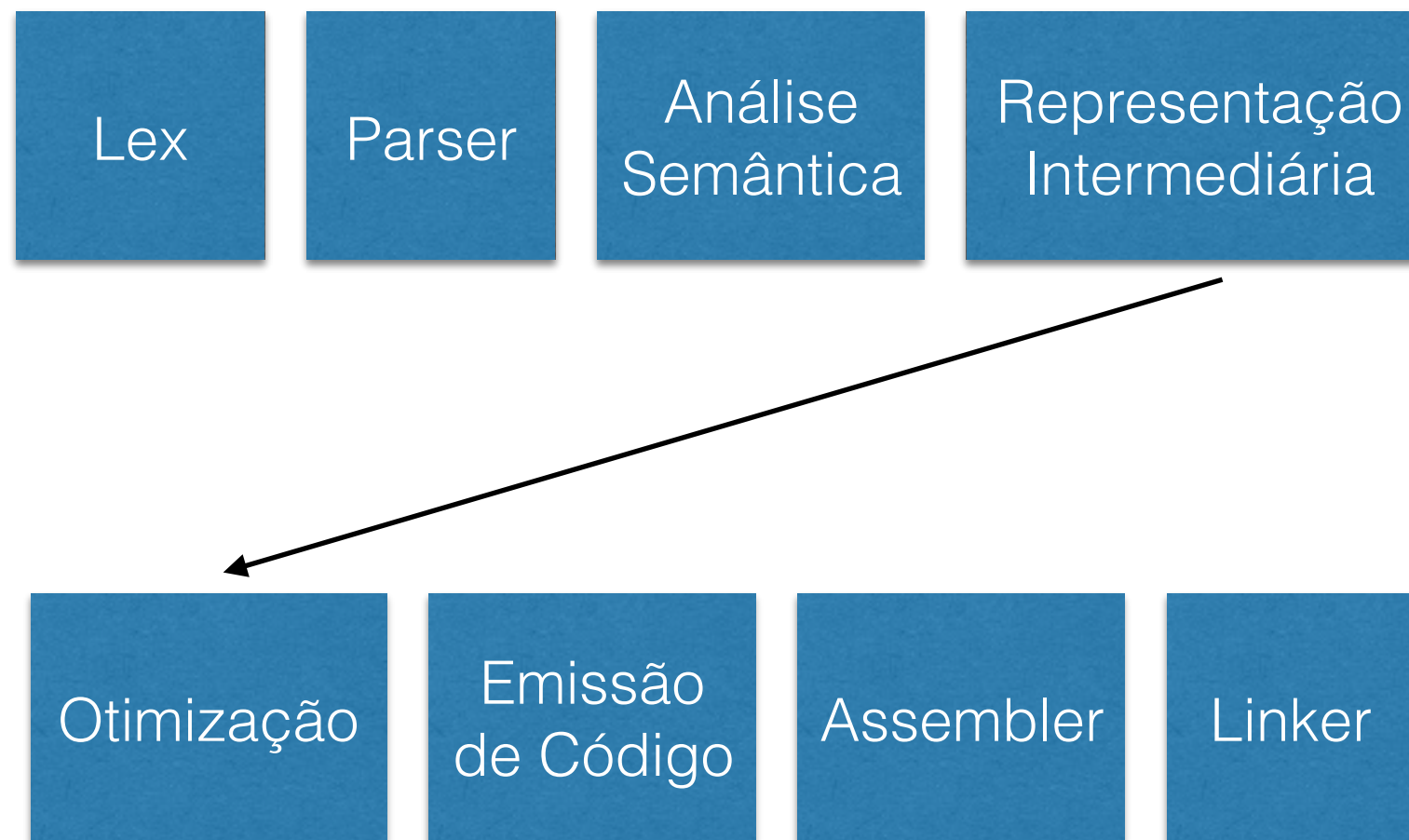
→

0101001001010
0100100010001
1101001000110

Exemplos de compiladores

Linguagem	Compilador	Resultado
C++	G++ / CLang	Assembly
Java	JAVAC	Bytecode
CoffeeScript	coffee	Javascript

Fases de um compilador



Análise léxica

- Quebra o arquivo fonte em palavras ou tokens

`int` `x` `=` `1` `+` `2` `*` `3`

Análise léxica

- Quebra o arquivo fonte em palavras ou tokens
- Expressões regulares
- Lex, Flex

Token	Expressão Regular	Exemplos
ID	[a-z][a-z0-9]*	foo bar123
NUM	[0-9]+	42 9
IF	if	if

Análise sintática

- Analisa a estrutura do program
- Gramática livre de contexto

$E \rightarrow E + E$

$E \rightarrow E * E$

$E \rightarrow \text{NUM}$

Análise sintática

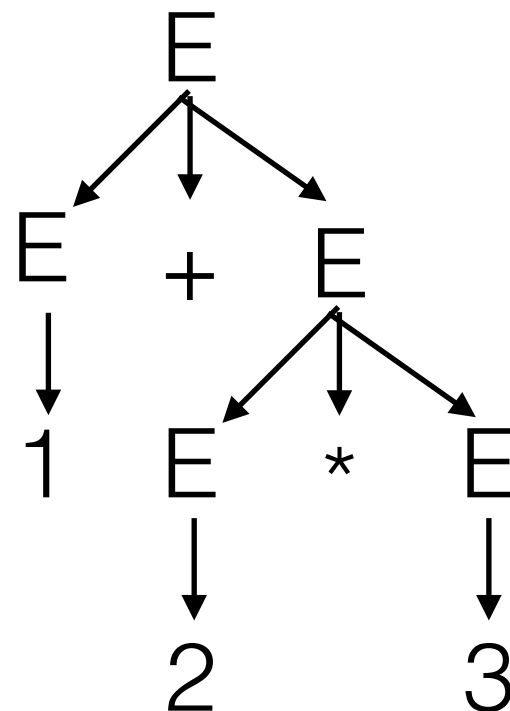
- Analisa a estrutura do program
- Gramática livre de contexto

$E \rightarrow E + E$

$E \rightarrow E * E$

$E \rightarrow \text{NUM}$

1 + 2 * 3

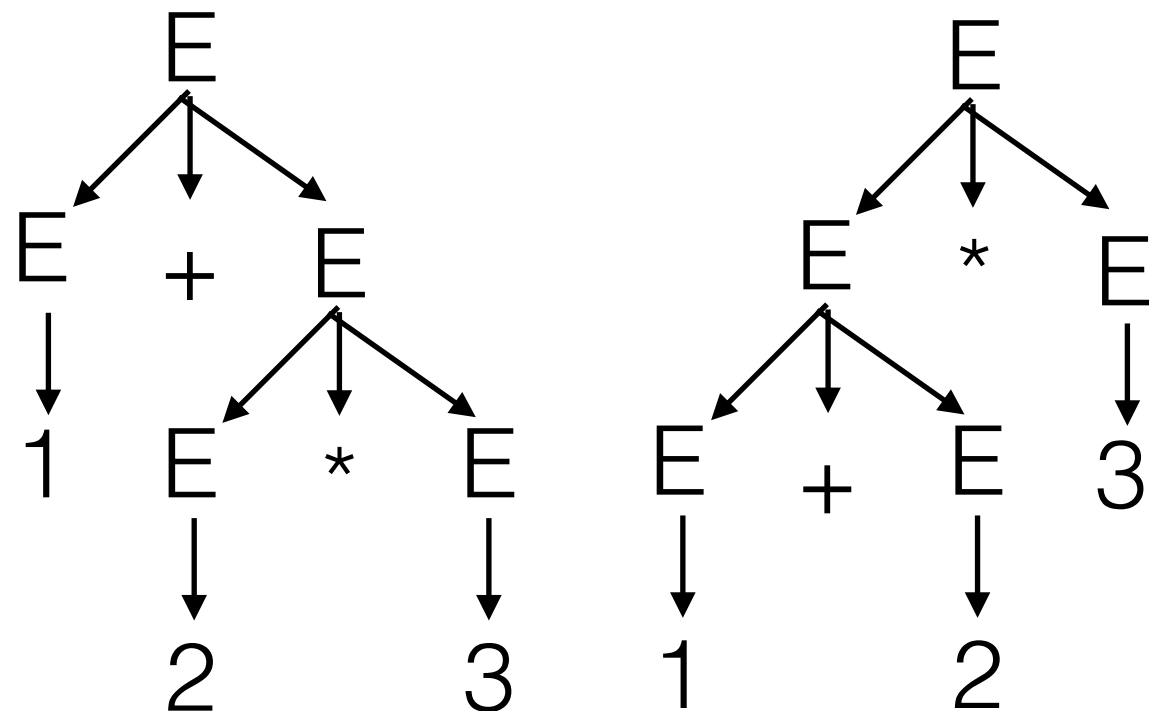


Análise sintática

- Analisa a estrutura do program
- Gramática livre de contexto

$E \rightarrow E + E$
 $E \rightarrow E * E$
 $E \rightarrow \text{NUM}$

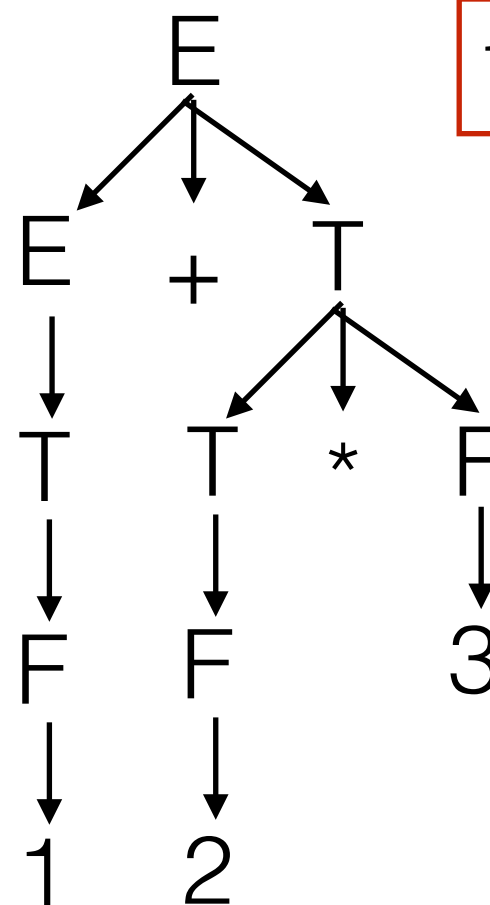
1 + 2 * 3



Análise sintática

- Analisa a estrutura do program
- Gramática livre de contexto

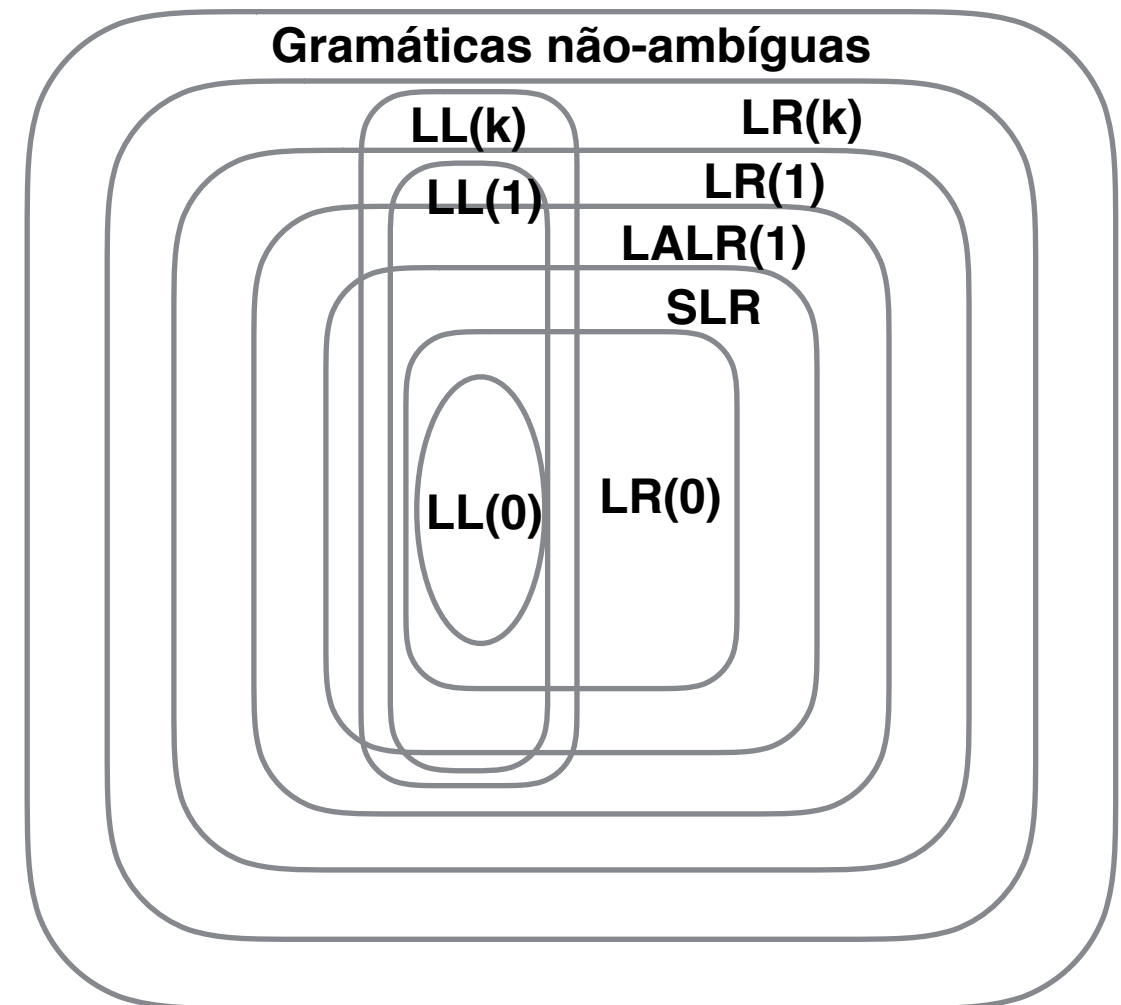
$E \rightarrow E + T$
 $E \rightarrow T$
 $T \rightarrow T * F$
 $T \rightarrow F$
 $F \rightarrow \text{NUM}$
 $F \rightarrow (F)$



1 + 2 * 3

Análise sintática

- Analisa a estrutura do program
- Gramática livre de contexto
- Yacc, Bison, Antlr



Parsers

- Compiladores comerciais normalmente utilizam um algoritmo recursivo descendente!
- GCC usava Bison... mas durante sua última refatoração o parser foi substituído por um recursivo descendente

Análise semântica

- Tabela de símbolos
- Verificação de tipos

```
int x = "blah"
```

```
int x = y
```

```
f(123)
```

```
obj.method(123)
```

Representação Intermediária

- Facilita diversas tarefas como
 - Otimização
 - Geração de código

Representação	Descrição	Exemplos
cálculo lambda	variável, definição de função e aplicação de função	Lisp, ML
Quádruplas	Corresponde à instruções de uma máquina de von Neumann	C, Java
SSA (Static single-assignment)	Valores são atribuídos a uma variável apenas uma vez	C, Java

Geração de código

- Código gerado
 - Assembly
 - Bytecode
 - Até mesmo javascript!

Geração de código

- LLVM
 - SSA (LLVM IR)
 - Implementa vários algoritmos de otimização
 - JIT
 - Gera código de máquina para várias plataformas

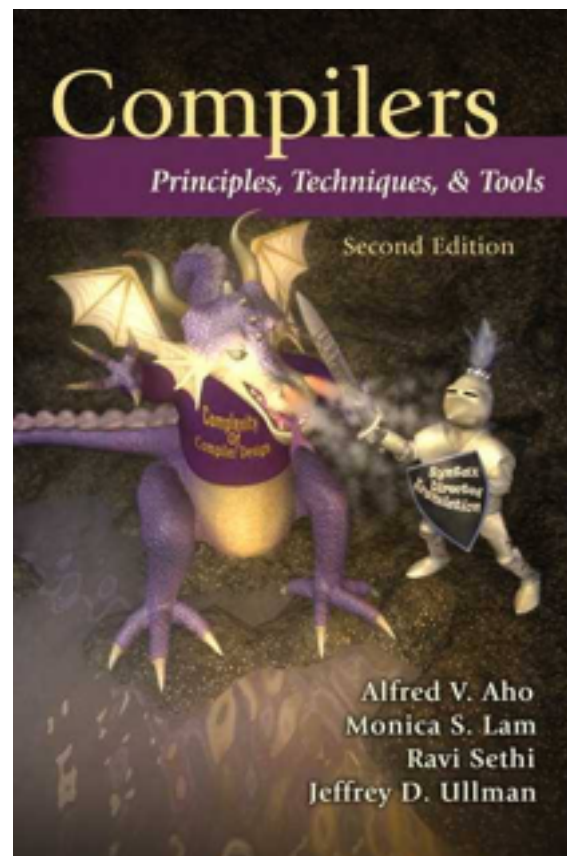
Linguagem	Compilador / Intepretador
C/C++	Clang
Ruby	Rubinius
Rust	rustc
Julia	julia

Open-source

Projeto	Linguagem	URL
Make you a Haskell	Haskell	http://dev.stephendiehl.com/fun/
Rubinius	Ruby	http://book.rubinius.com
CoffeeScript	CoffeeScript	https://github.com/jashkenas/coffeescript

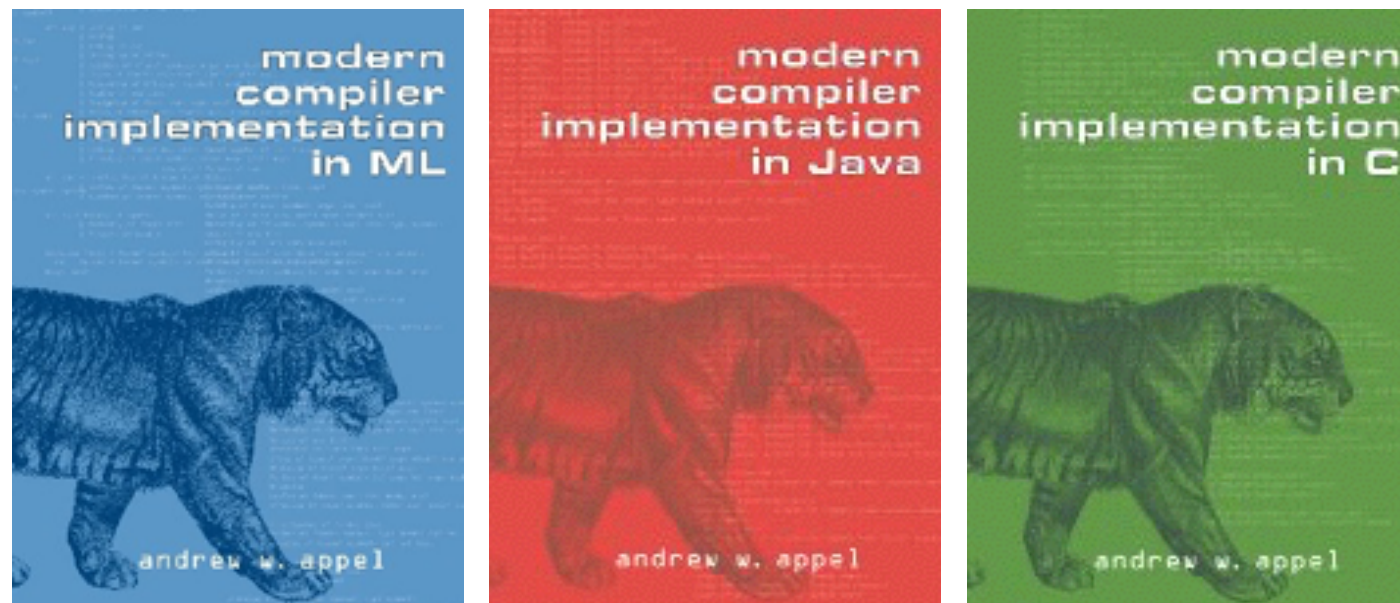
Livros

- Compiler: Principles, Techniques & Tools
Alfred V. Aho, Monica S. Lam, Ravi Sethi, and
Jeffrey D. Ullman



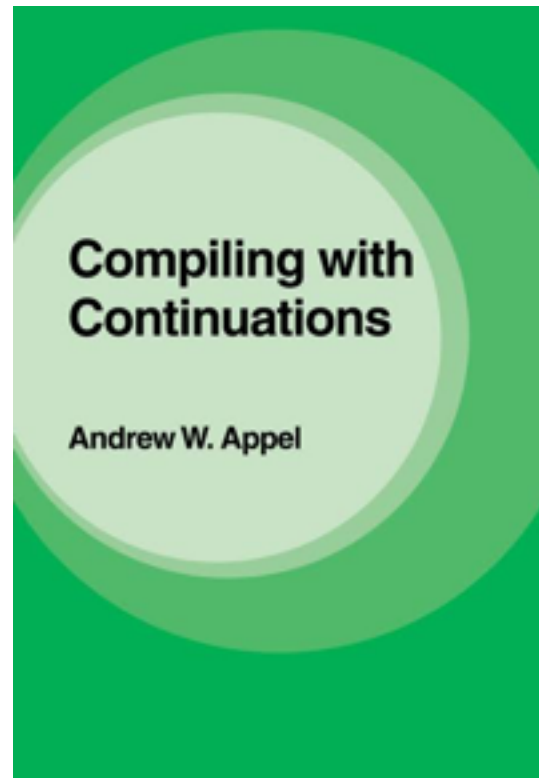
Livros

- Modern Compiler Implementation in X
Andrew W. Appel



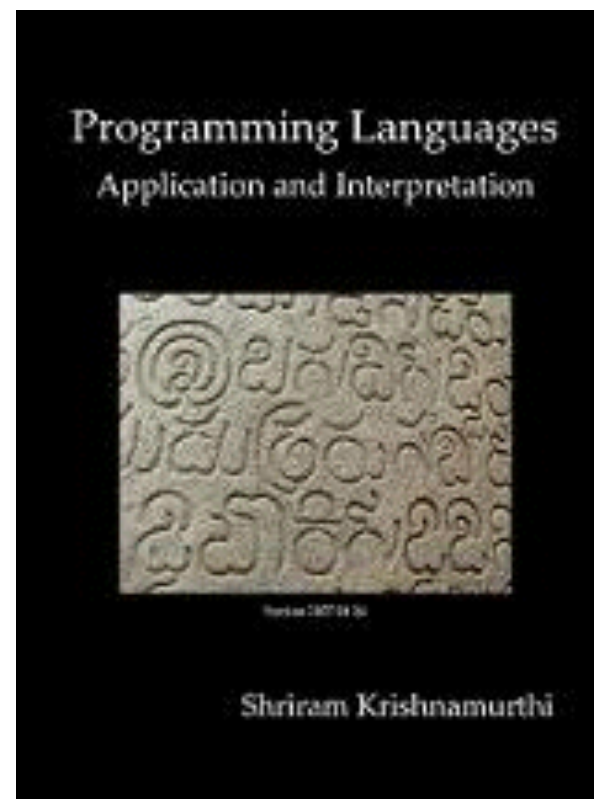
Livros

- Compiling with Continuations
Andrew W. Appel



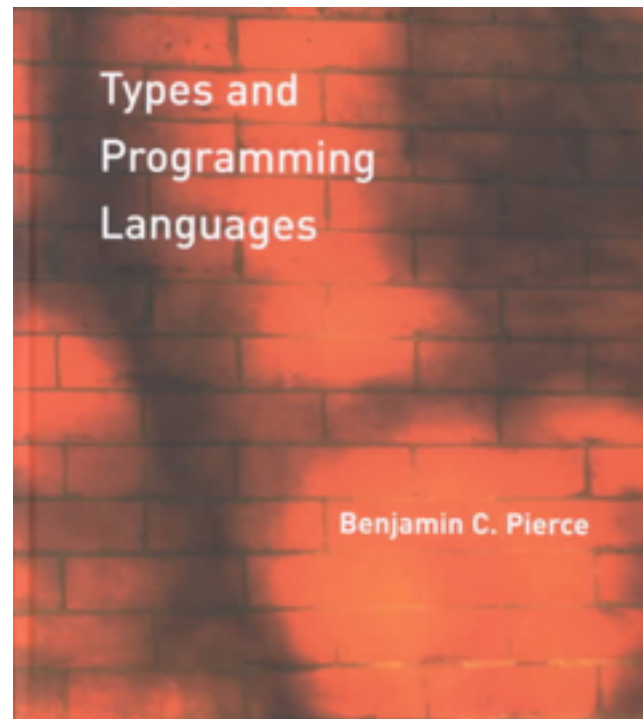
Livros

- Programming Languages: Application and Interpretation
Shriram Krishnamurthi



Livros

- Types and Programming Languages
Benjamin C. Pierce



Obrigado!