

GOPHER CHINA 2020

中国 上海 / 2020-11.21-22

Go+ 演进之路

姜智@七牛云



大纲

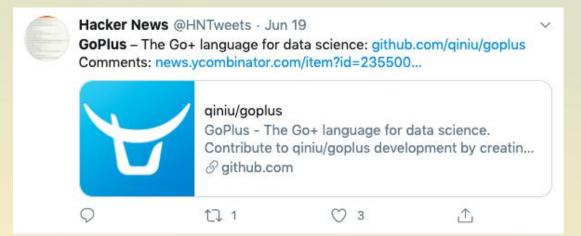
• Go+ 发展历程

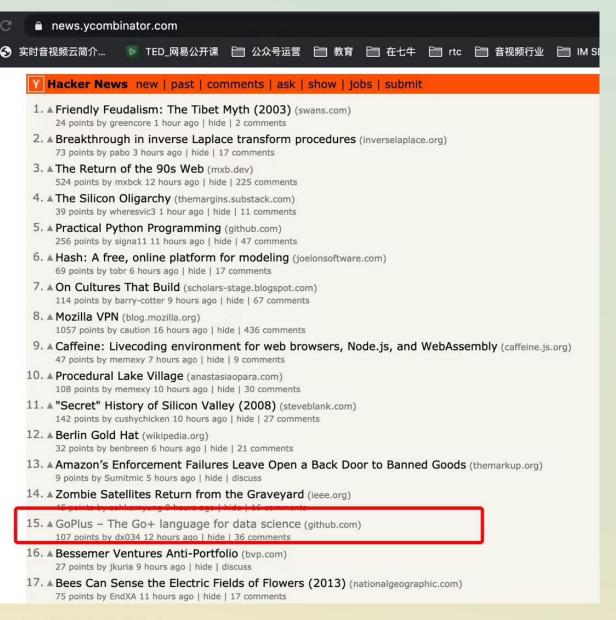
• Go+ 特性

• Go+ 实现解析

Go+ 意外走红

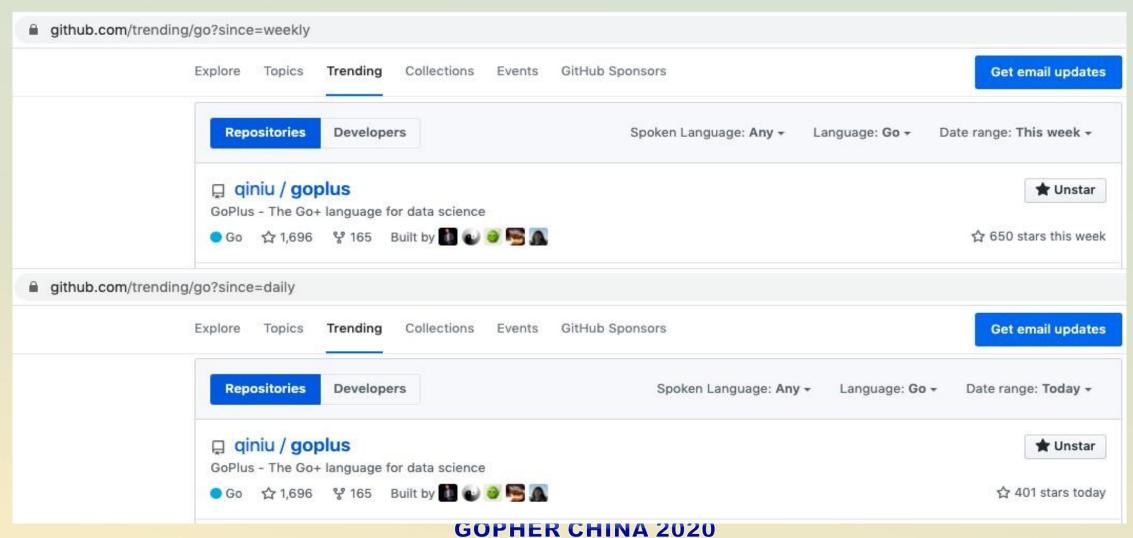






GOPHER CHINA 2020

Go+ 意外走红



中国 上海 / 2020-11.21-22



mattn

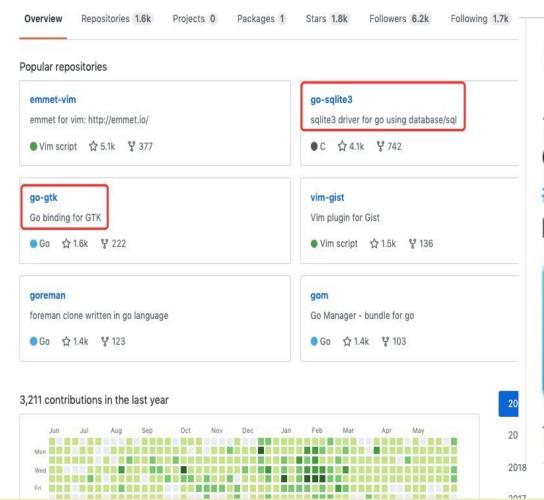
mattn

O Sponsor

Unfollow

Long-time Golang user&contributor, Google Dev Expert for Go, and author of many Go tools, Vim plugin author. Windows hacker C#/Java/C/C++

- Osaka, Japan
- mattn.jp@gmail.com
- https://mattn.kaoriya.net/
- mattn_ip





データサイエンスに向いた Go 言語の拡張言語。ほぼ Go 言語のまま型を書かなくて良いスクリプト言語。 #golang / "GitHub - qiniu/goplus: GoPlus - The Go+ language for data science"

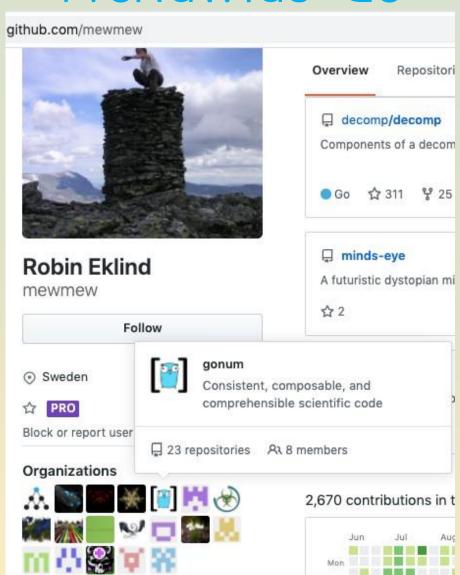


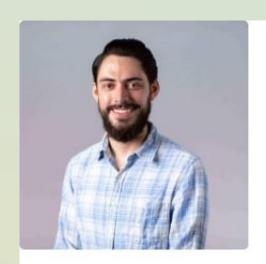
qiniu/goplus

GoPlus - The Go+ language for data science. Contribute to qiniu/goplus development by creating an account on GitHub. Sighthub.com

下午2:43 - 2020年6月9日 - はてなブックマーク

GOPHER CHINA 2020





Nate Fischer nfischer

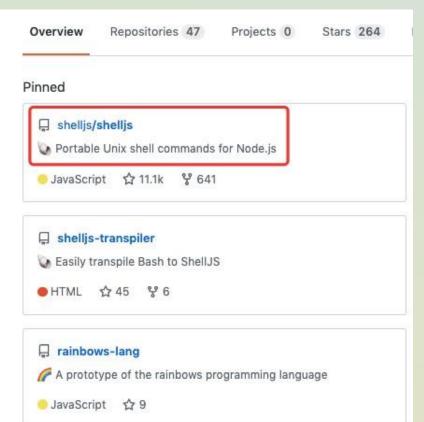
Unfollow

@google engineer working on Chromium and Android WebView. @shelljs lead dev

- @google
- Mountain View, California
- https://nfischer.github.io/

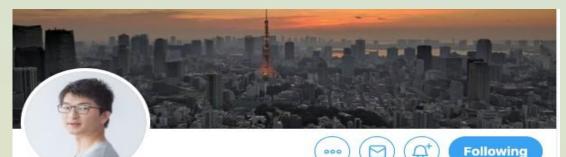
Block or report user

Organizations



145 contributions in the last year





Yoshi Yamaguchi

@ymotongpoo

Tweets

Developer Advocate of @Google Cloud for Observability & Go; Google Cloud Operations and OpenTelemetry. An enthusiastic Gopher. Tweets are on my own.

◎ 日本 東京 Ⅲ Joined April 2007

654 Following 8,378 Followers

Followed by ドッグ and mattn

Yoshi Yamaguchi
Qymotongpoo · 47m

Tweets & replies

I have heard this programming language name as some sort of joke but it became reality. ainiu/goplus: GoPlus - The Go+ language for data science



giniu/goplus

GoPlus - The Go+ language for data science. Contribute to giniu/goplus development by creatin... @github.com

Media







Likes





or jump to...

Pull requests Is



Misha Brukman

mbrukman

Follow

Co-Founder, @JanusGraph • Startup/VC advisor

@google

A https://misha.brukman.net/

Block or report user

Organizations











Trishank Karthik Kuppusamy

trishankatdatadog

Follow

Staff Security Engineer at @DataDog. Helped to research and develop @theupdateframework and @uptane.

(@DataDog

New York, NY

A https://keybase.io/trishankdatadog

Block or report user

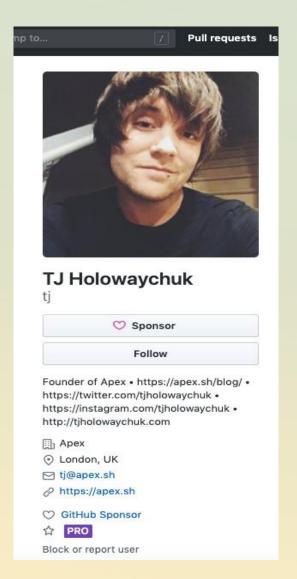
Organizations

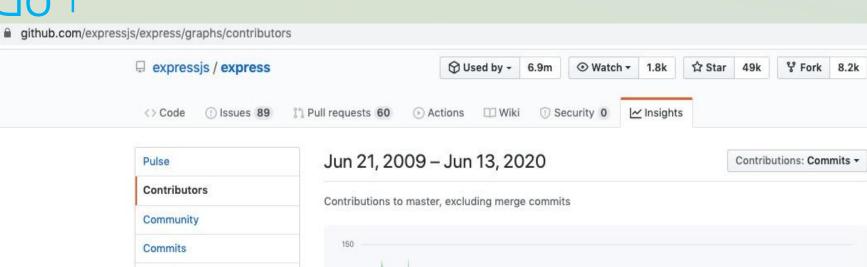














2013

100

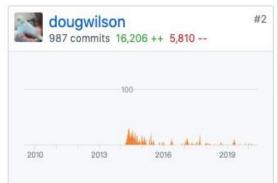
50

Code frequency

Network

Forks

Dependency graph



Why Go+

• 正面: 为什么这么多牛人会关注 Go+?

• 反面:好像很难搞的样子,为什么要干?

• 反面: Python 这么强, 怎么打得过?

Data Science 的发展

- 从前
 - Limited Domains (有限领域): 比如 BI (Business Intelligence)
 - Limited Data (有限数据规模): 比如 Excel、Matlab
- 未来
 - Full Domains (全领域): 智能应用 (Intelligent Application)
 - 典型代表: 抖音、快手
 - Big Data (大规模数据)
 - Any Where (随处): 云 (Cloud)、智能手机 (Small Phone)、嵌入式设备 (IoT)

中国 上海 / 2020-11 21-22

信息科技发展的自然结果

- 应用极大化地丰富(充分竞争)
 - 差异化竞争让应用越来越"聪明"— Intelligent Application
- 数字化信息(数据)极大化地产生

这就是 DT 时代

• IT => DT

- 数据地位的变化
 - Businesses Intelligence => Intelligent Application
 - •数据是副产品 => 数据是原材料(石油),无处不在,深植于业务流
- Data Science 的地位变化
 - 这意味着, Data Science 将基础设施化
 - 数学软件 (Application) => 基础设施 (Infrastructure)

GOPHER CHINA 2020

数据科学领域 技术栈

Business Intelligence (商业智能) Math DSL (数学语言: M, etc.) Application (应用软件) Mathematic Software (数学软件: Matlab, etc.) Script (脚本语言: Python, etc.) General Language (通用语言: Go, Java, etc.) OS User Interface (操作系统使用界面) System Language (系统级语言: C, Go, etc.) OS Core (操作系统内核) Bare-metal based Language (裸金属语言: ASM, C, etc.) CPU / FPU (中央处理器 / 浮点协处理器) OUT TIEN CITINA ZUZU

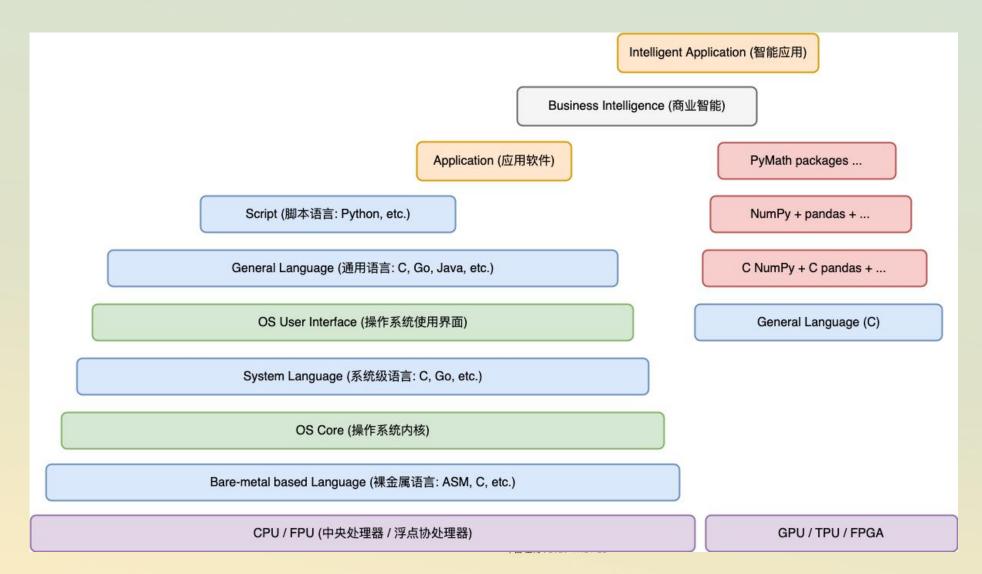
变化难度

- 经典的汉诺塔问题 (假设只改变一个依赖)
 - f(n) = 2*f(n-1) + 1
 - f(1) = 1
- 改变技术依赖栈是极其困难的

数据科学的汉诺塔第一层迁移

- 数学软件 Python 化
 - 数学软件的平民化(与脚本语言结合)
- 汉诺塔定则: 只能一层层迁移
 - 技术栈迁移的步子迈得越大,越不可完成!

技术依赖栈



Python 不会是 Data Science 的终局

• 因为, Python 成不了基础设施(Infrastructure)。

• Data Science 本质上是算力革命,是计算密集型的业务。

• Data Science 进一步下沉,终局会是什么?

Go+ for Data Science

- 终局是: 数学软件与通用语言的融合
 - 完成 Data Science 基础设施化
- 所以 Go+ 来了!

汉诺塔的再一次迁移

Intelligent Application (智能应用) Script (脚本语言: Python, etc.) General Language (通用语言: C, Go, Java, etc.) GopMath packages ... OS User Interface (操作系统使用界面) NumGoPlus + pandas + ... System Language (系统级语言: C, Go, etc.) Go+ General Language (Go) OS Core (操作系统内核) Bare-metal based Language (裸金属语言: ASM, C, etc.) CPU / FPU (中央处理器 / 浮点协处理器) GPU / TPU / FPGA

GOPHER CHINA 2020

中国 上海 / 2020-11.21-22

大纲

• Go+ 发展历程

• Go+ 特性

• Go+ 实现解析

Go+ 特性

• 静态语言

• 与Go完全兼容

• 语法简洁(for data science)

Go+与 Go的互操作

• Go+ 将支持所有Go的feature(已经支持 基本语法,流程控制、结构体方法、defer, goroutine, channel等)

• 所有Go package均可以被Go+ import

• 所有Go+的package均可以被convert成Go package, 并被Go来import

Go+双引擎

• Bytecode backend

• Go code generation

Go+ 示例

For example, the following is legal Go+ source code:

```
a := [1, 2, 3.4]
println(a)
```

How do we do this in the Go language?

```
package main

func main() {
    a := []float64{1, 2, 3.4}
    println(a)
}
```

运行方式

Bytecode backend

Go code generation

\$ gop go var_and_op.gop

```
$ cat var_and_op.gop
#!/usr/bin/env gop run
x := 123.1 - 3i
y, z := "Hello, ", 123
println(y+"complex:", x+1, "int:", z)
 in ~/goplus/tutorial/02-Var-
 gop run var_and_op.gop
Hello, complex: (124.1-3i) int: 123
 in ~/goplus/tutorial/02-Var-
 ./var_and_op.gop
Hello, complex: (124.1-3i) int: 123
```

Go+ 相关

- Repl
- Go+ playground
- Go+ vscode plugin

```
play.goplus.org
$ gop
Gop is a tool for managing Go+ source code.
                                                                                    Run Format Share
                                                                                                          Hello, Go+
                                                                                                                                 (7) Star
                                                           he Go+ Playground
Usage:
                                                            1 println("Hello, Go+")
                                                            3 println(1r << 129)</pre>
        gop <command> [arguments]
                                                            4 println(1/3r + 2/7r*2)
The commands are:
                                                            6 arr := [1, 3, 5, 7, 11, 13, 17, 19]
                                                             println(arr)
                                                            8 \text{ println}([x*x \text{ for } x \leftarrow arr, x > 3])
                    Run a Go+ program
        run
                    Convert Go+ packages into Go packages
        qo
                                                             m := {"Hi": 1, "Go+": 2}
        fmt
                    Format Go+ packages
                                                           11 println(m)
                    Export Go packages for Go+ programs
        export
                                                             println({v: k for k, v <- m})</pre>
                                                           repl
                    Play Go+ in console
                                                          14 println([v for v <- m])
                                                                中国 上海 / 2020-11.21-22
```

Go+ 特有语法

• 有理数

Map

```
x := {"Hello": 1, "xsw": 3.4} // map[string]float64
y := {"Hello": 1, "xsw": "Go+"} // map[string]interface{}
z := {"Hello": 1, "xsw": 3} // map[string]int
empty := {} // map[string]interface{}
```

Slice

```
x := [1, 3.4] // []float64
y := [1] // []int
z := [1+2i, "xsw"] // []interface{}
a := [1, 3.4, 3+4i] // []complex128
b := [5+6i] // []complex128
c := ["xsw", 3] // []interface{}
empty := [] // []interface{}
```

Go+ 特有语法

 List/map comprehension

```
a := [x*x for x <- [1, 3, 5, 7, 11]]
b := [x*x for x <- [1, 3, 5, 7, 11], x > 3]
c := [i+v for i, v <- [1, 3, 5, 7, 11], i%2 == 1]
d := [k+","+s for k, s <- {"Hello": "xsw", "Hi": "Go+"}]

arr := [1, 2, 3, 4, 5, 6]
e := [[a, b] for a <- arr, a < b for b <- arr, b > 2]

x := {x: i for i, x <- [1, 3, 5, 7, 11]}
y := {x: i for i, x <- [1, 3, 5, 7, 11], i%2 == 1}
z := {v: k for k, v <- {1: "Hello", 3: "Hi", 5: "xsw", 7: "Go+"}, k > 3}
```

For range

```
sum := 0
for x <- [1, 3, 5, 7, 11, 13, 17], x > 3 {
    sum += x
}
```

Go+ 特有语法

```
Error handling
                                               import (
                                                 "strconv"
expr! // panic if err
expr? // return if err
                                               func add(x, y string) (int, error) {
expr?:defval // use defval if err
                                               return strconv.Atoi(x)? + strconv.Atoi(y)?, nil
                                               func addSafe(x, y string) int {
                                                 return strconv.Atoi(x)?:0 + strconv.Atoi(y)?:0
                                               println(~add("100", "23"):~, add("100", "23")!)
                                               // add("100", "23"): 123
                                               println( addSafe("10", "abc"): , addSafe("10", "abc"))
                                     GOPHER (// addSafe("10", "abc"): 10
```

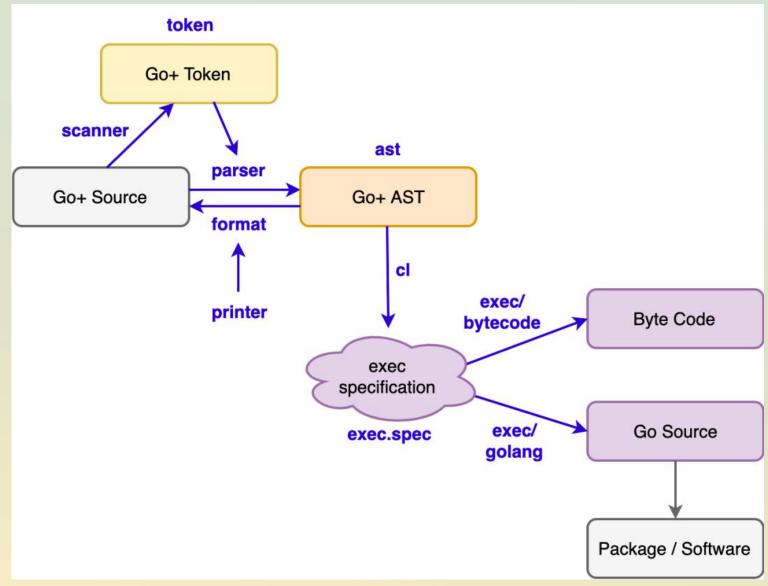
大纲

• Go+ 发展历程

• Go+ 特性

• Go+ 实现解析

Go+ 运行逻辑



Go+ 目录结构

- github.com/goplus/gop/token
- github.com/goplus/gop/scanner
- github.com/goplus/gop/parser
- github.com/goplus/gop/ast
- github.com/goplus/gop/format
- github.com/goplus/gop/printer
- github.com/goplus/gop/cl
- github.com/goplus/gop/exec.spec
- github.com/goplus/gop/exec/bytecode
- github.com/goplus/gop/exec/golang

Go+ Token

- 词法分析:将Go+原文件转换成 Token序列
- 主要实现: github.com/goplus/gop/scanner
- Go+ 独有token:
 - Error handling:?
 - 有理数: r

```
1 y, z := "Hello, ", 123
2 println("string:", y, "int:", z)
```



Token

```
文件:行号:列号
                Token名
                          Token面值
hello.go:2:1
                        "y"
                IDENT
                        1111
hello.go:2:2
                        H_2H
hello.go:2:4
                IDENT
hello.go:2:6
                :=
                        "\"Hello, \""
hello.go:2:9
                STRING
hello.go:2:18
hello.go:2:20
                        "123"
                INT
                        "\n"
hello.go:2:23
hello.go:3:1
                        "println"
                IDENT
hello.go:3:8
hello.go:3:9
                        "\"string:\""
                STRING
hello.go:3:18
                        "y"
hello.go:3:20
                IDENT
hello.go:3:21
hello.go:3:23
                        "\"int:\""
                STRING
hello.go:3:29
                        11.11
                        "7"
hello.go:3:31
                IDENT
hello.go:3:32
hello.go:3:33
                         "\n"
```

Go+ AST

- 语法分析:将Token序列转换为 抽象语法树(AST)
- 主要实现:
 github.com/goplus/gop/parser
- Go+和Go的抽象语法树的区别?

```
1  y, z := "Hello, ", 123
2  println("string:", y, "int:", z)
```

```
24 . . . Body: *ast.BlockStmt {
25 . . . Lbrace: 1:26
26 . . . List: []ast.Stmt (len = 2) {
27 . . . . 0: *ast.AssignStmt {
28 . . . . Lhs: []ast.Expr (len = 2) {
  . . . . . . . 0: *ast.Ident {
30 . . . . . . NamePos: 1:27
31 . . . . . . Name: "y"
32 . . . . . . . . Obj: *ast.Object {
33 . . . . . . . . Kind: var
34 . . . . . . . Name: "y"
  . . . . . . . . Decl: *(obj @ 27)
  . . . . . . . 1: *ast.Ident {
  . . . . . . . NamePos: 1:30
40 . . . . . . Name: "z"
  . . . . . . . . Obj: *ast.Object {
42 . . . . . . . . Kind: var
43 . . . . . . . Name: "z"
  . . . . . . . . Decl: *(obj @ 27)
. . . . . . . . .
  . . . . . TokPos: 1:32
   . . . . . Tok: :=
  . . . . . . Rhs: []ast.Expr (len = 2) {
51 . . . . . . 0: *ast.BasicLit {
  . . . . . . . . ValuePos: 1:35
  . . . . . . . Kind: STRING
  . . . . . . . . Value: "\"Hello, \""
   . . . . . . . 1: *ast.BasicLit {
   . . . . . . . . ValuePos: 1:46
                     Kind: TNT
```

Go+ AST

- Go+ 独有语法
- •需要Go+的AST实现

```
1 x := [1, 3.4]
2 println(x)
```

```
24 . . . Body: *ast.BlockStmt {
25 . . . Lbrace: 1:26
26 . . . List: []ast.Stmt (len = 2) {
27 . . . . 0: *ast.AssignStmt {
28 . . . . . Lhs: []ast.Expr (len = 1) {
29 . . . . . . 0: *ast.Ident {
30 . . . . . . NamePos: 1:27
31 . . . . . . Name: "x"
32 . . . . . . . . . . . . Obj: *ast.Object {
33 . . . . . . . Kind: var
35 . . . . . . . Decl: *(obj @ 27)
36 . . . . . . . . . .
37 . . . . . . . . .
38 . . . . . . . .
39 . . . . . TokPos: 1:29
40 . . . . . Tok: :=
41 . . . . . . Rhs: []ast.Expr (len = 1) {
42 . . . . . . 0: *ast.SliceLit {
43 . . . Lbrack: 1:32
44 . . . . . . . . Elts: []ast.Expr (len = 2) {
45 . . . . . . . . 0: *ast.BasicLit {
46 . . . . . . . . ValuePos: 1:33
47 . . . . . . . . Kind: INT
50 . . . . . . . . 1: *ast.BasicLit {
51 . . . . . . . . ValuePos: 1:36
52 . . . . . . . . Kind: FLOAT
53 . . . . . . . . Value: "3.4"
56 . . . . . . Rbrack: 1:39
57 . . . . . . . Incomplete: false
58 . . . . . . . . .
59 . . . . . . .
```

Go+ ASI

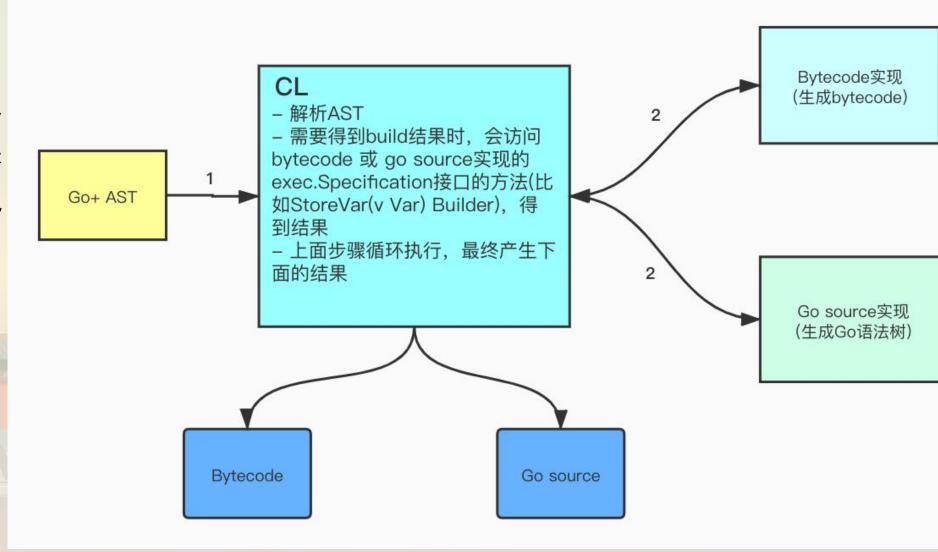
- Go+ SliceLit
- 抽象语法树实现

```
1 x := [1, 3.4]
2 println(x)
```

```
func (p *parser) parseArrayType() ast.Expr {
                                                                                                              func (p *parser) parseArrayTypeOrSliceLit(allowSliceLit bool) (expr ast.Expr, isSliceLit boo
  if p.trace {
                                                                                                                 if p.trace {
       defer un(trace(p, "ArrayType"))
                                                                                                                     defer un(trace(p, "ArrayType"))
   lbrack := p.expect(token.LBRACK)
                                                                                                                 lbrack := p.expect(token.LBRACK)
   p.exprLev++
                                                                                                                 p.exprLev++
   var len ast.Expr
                                                                                                                 var len ast.Expr
   if p.tok == token.ELLIPSIS {
                                                                                                                 if p.tok == token.ELLIPSIS {
       len = &ast.Ellipsis{Ellipsis: p.pos}
                                                                                                                     len = &ast.Ellipsis{Ellipsis: p.pos}
      p.next()
                                                                                                                     p.next()
   } else if p.tok != token.RBRACK {
                                                                                                                 } else if p.tok != token.RBRACK {
       len = p.parseRHS()
                                                                                                                     len = p.parseRHS()
                                                                                                                     if allowSliceLit && p.tok == token.COMMA { // [a, b, c, d ...]
   p.exprLev--
                                                                                                                         elts := p.parseSliceLit(lbrack, len)
   p.expect(token.RBRACK)
   elt := p.parseType()
   return &ast.ArrayType{Lbrack: lbrack, Len: len, Elt: elt}
                                                                                                                 rbrack := p.expect(token.RBRACK)
unc (p *parser) makeIdentList(list []ast.Expr) []*ast.Ident {
   idents := make([]*ast.Ident, len(list))
                                                                                                                 var elt ast.Expr
   for i, x := range list {
                                                                                                                 if allowSliceLit {
                                                                                                                     elt = p.tryType()
       ident, isIdent := x.(*ast.Ident)
      if !isIdent {
           if _, isBad := x.(*ast.BadExpr); !isBad {
               // only report error if it's a new one
               p.errorExpected(x.Pos(), "identifier")
                                                                                                                 } else {
                                                                                                                     elt = p.parseType()
           ident = &ast.Ident{NamePos: x.Pos(), Name: "_"}
                                                                                                                 if p.trace {
                                                                                                                     log.Debug("parseArrayType:", len, "elt:", elt)
   return idents
```

Go+ 编译和双擎

- Go+双擎: Bytecode & Go Source
- Executing Specification, 定义双擎输出编译结果的接口,双擎独立实现各自的编译输出行为。



Go+ ByteCode

- 编译 针对Go+ AST编译生成bytecode 编译代码:
 - 发现是ast . AssignStmt
 - -拿2个右值ast.Rhs,分别压栈
 - -轮询左值ast.Lhs,将右值出栈,校验
 - 生成byte code

```
1  y, z := "Hello, ", 123
2  println("string:", y, "int:", z)
```

bytecode

右移26bit为操作指令,低位记录数据偏移等

Go+ ByteCode

• 执行阶段,读取bytecode和数据进行操作

```
1 y, z := "Hello, ", 123
2 println("string:", y, "int:", z)
```

bytecode执行过程

bytecode	操作指令	数据栈
110000000000000000000000000000000000000	pushConstR	D
11000000000000000001111011	pushInt	[Hello,]
100000000000000000000000000000000000000	storeVar	[Hello, 123]
100000000000000000000000000000000000000	storeVar	[Hello,]
110000000000000000000000000000000000000	pushConstR	D
111100000000000000000000000000000000000	loadVar	[string:]
1100000000000000000000000000101	pushConstR	[string: Hello,]
111100000000000000000000000000000000000	loadVar	[string: Hello, int:]
10000000100000000000000011	callGoFuncv	[string: Hello, int: 123]

GOPHER CHINA 2020

Go+ 生成Go Source

- 编译阶段,将Go+ AST转换为Go AST
 - 轮询Go+ AST语句,将表达式入栈
 - 当需要生成一个Go AST的语句时,出栈数据进行组装
- •运行阶段,从Go AST转换为Go source

Go+ 下一步

- 当前
 - Go+的语法
 - 支持Go的feature
- 未来, 专注数据科学领域
 - GoTorch
 - Numgoplus
 - GopMath
 - Pandas

. . .

• 把冷板凳捂热,让Go+助力数据科学



GOPHER CHINA 2020

中国 上海 / 2020-11.21-22

Thanks

https://github.com/goplus/gop

