



Rust 编程语言入门



Microsoft®
Most Valuable
Professional

杨旭，微软MVP
Rust、Go、C#开发者

8.4 String (下)

对 String 按索引的形式进行访问

- 按索引语法访问 String 的某部分，会报错（例子）
- Rust 的字符串不支持索引语法访问

内部表示

- String 是对 Vec<u8> 的包装
 - len() 方法
- （例子）

字节、标量值、字形簇

Bytes, Scalar Values, Grapheme Clusters

- Rust 有三种看待字符串的方式：
 - 字节
 - 标量值
 - 字形簇（最接近所谓的“字母”）
- （例子）
- Rust 不允许对 String 进行索引的最后一个原因：
 - 索引操作应消耗一个常量时间($O(1)$)
 - 而 String 无法保证：需要遍历所有内容，来确定有多少个合法的字符。

切割 String

- 可以使用 **[]** 和 **一个范围** 来创建字符串的切片（例子）
 - 必须谨慎使用
 - 如果切割时跨越了字符边界，程序就会 panic。
 - **(b1,b2),(b3,b3),(b4,b5),(b7,b7)**
 - panic

遍历 String 的方法

- 对于标量值：chars() 方法
 - （例子）
- 对于字节：bytes() 方法：
 - （例子）
- 对于字形簇：很复杂，标准库未提供。

String 不简单

- Rust 选择将正确处理 String 数据作为所有 Rust 程序的默认行为
 - 程序员必须在处理UTF-8数据之前投入更多的精力
- 可防止在开发后期处理涉及非 ASCII 字符的错误。

再见

