

1 Rust语言介绍

Rust语言其他方面无需多言，只要强调一点就是它的目标：性能、安全以及实用

1.1 内存安全方案

1.1.1 Rust针对C语言的不足

1. 禁止对空指针和悬垂指针解引用

空指针指的是指向了不存在的数据

悬垂指针指的是原本指向的数据被释放掉了

2. 读取未初始化的内存
3. 非法释放已经释放或未分配的指针
4. 缓冲区溢出

1.1.2 安全无缝地沟通C语言

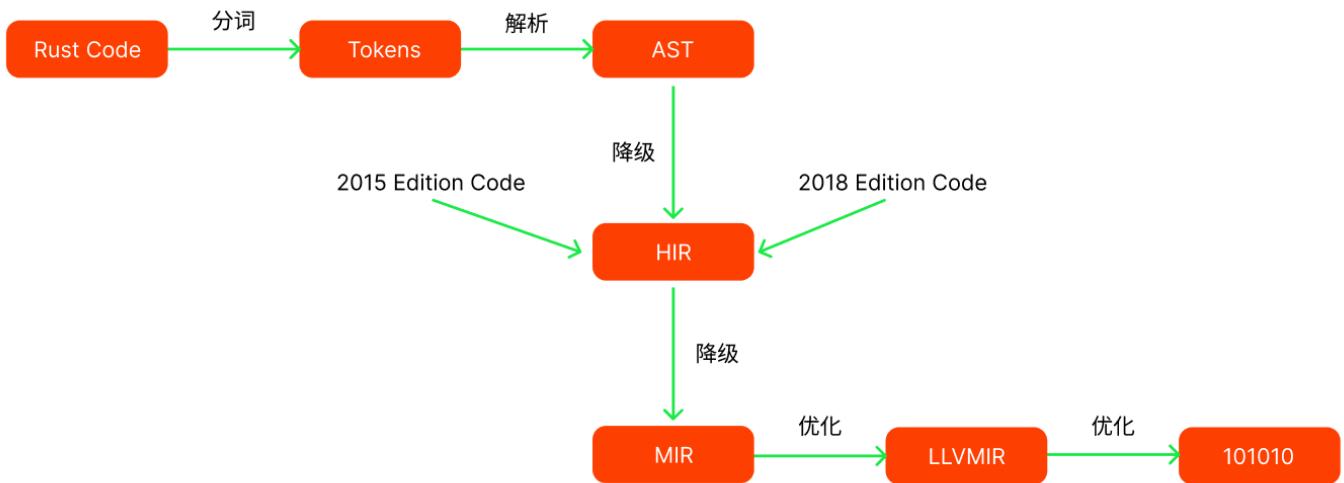
通过C-ABI零成本和C语言打交道

划分了Safe Rust和Unsafe Rust

2 Rust语言基础

2.1 Rust 语言编译

2.1.1 编译过程



2.1.2 Rust与其它语言编译比较

大部分语言会将词条流解析到的抽象语法树直接转为机器码，但是rust会将其转为高级中间语言以及中级中间语言、LLVM中间语言，然后再交由LLVM后端生成机器码。各级中间语言承担的功能：

1. 高级中间语言：类型检查、方法查找
2. 中级中间语言：借用检查、代码生成、泛型单态化、优化等工作
3. Rust语言版次差异在到达中级中间语言时就会消除

2.2 Rust 词法结构

词法结构对于任何一种语言来说都非常重要，因为它不光是构成语言的必要部分，而且也关乎到语言如何解析和编译。在rust中，词法结构中的词条还涉及元编程

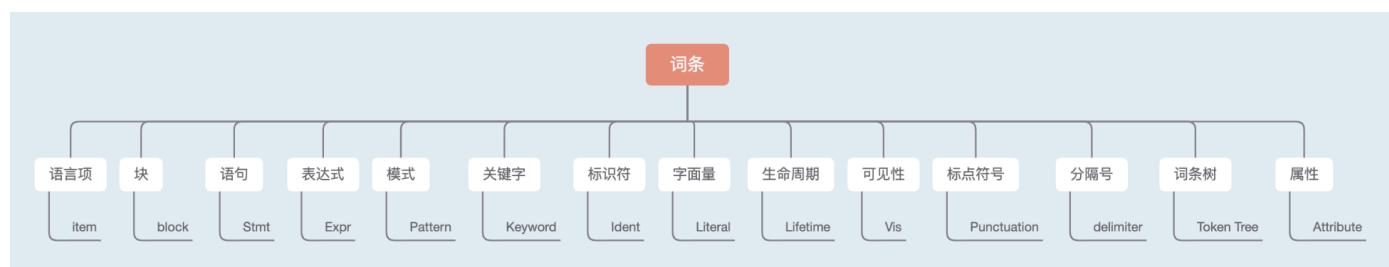
2.2.1 六大词法结构

1. 关键字：严格关键字、保留字、弱关键字
2. 标识符：不以数字开头的ASCII字符注释

```
let name = "name";
let _100 = "number";
let math_grade = 150;

println!("{} , {} , {}" , name , _100 , math_grade)
```

3. 注释：Rust可以使用注释直接生成文档，非常友好。
4. 空白：空白不表示任何含义，如换行等
5. 词条：词条在写宏的时候非常有用（它是宏的关键词，需要熟悉并深刻理解词条才能编写宏代码），Rust语言有14个词条



6. 路径：Rust中路径有三种用途，模块路径、方法调用和泛型类型指定

```
pub fn main() {

    /// 1.模块路径
    ///
    pub mod a {
        fn foo() {
            println!("a")
        }
    }
}
```

```
}

pub mod b {
    pub mod c {
        pub fn foo() {
            super::super::foo();
            self::super::super::foo();
        }
    }
}

a::b::c::foo();

/// 2.方法调用
///
struct S;

impl S {
    fn correlation_function(){
        println!("correlation function");
    }
}

trait T1 {
    fn method1() {
        println!("method1");
    }
}

impl T1 for S {}

trait T2 {
    fn method2() {
        println!("method2")
    }
}

impl T2 for S {}

// 注意: 调用方法有两种情况
// 两个trait中的方法相同时使用完全限定无歧义调用
<S as T1>::method1();
<S as T2>::method2();

// 其他情况下, 调用关联函数和方法的方式相同
S::correlation_function();
S::method1();

/// 3.泛型函数-turbofish操作符
```

```
///  
  
// 将0到9收集到vec中，默认类型是i32，但是可以指定为u64  
let vec0 = (0..10).collect::<Vec<_>>();  
let vec1 = (0..10).collect::<Vec<u64>>();  
println!("{:?}", vec1);  
  
// 开辟一个容量为1024的u8Vec  
let vec2 = Vec::<u8>::with_capacity(1024);  
  
println!("{:?}", vec2);  
}
```

2.3 Rust 语法骨架

Rust语法骨架只包含三类元素

1. 属性：行属性和块属性

以# 或者 #! 开头

2. 分号：行分隔符

以 ; 结尾

3. 花括号：块分隔符

以 } 结尾

2.4 Rust表达式

在Rust中，一切皆表达式，它是以分号 ; 和花括号 {} 进行区分，而不是以循环、匹配等条件作为区分

一切皆表达式可以引申为一切皆类型，因为表达式都有值，而值都有类型

2.4.1 表达式分类：按语法骨架

其中作为Rust骨架的分号和花括号构成了Rust语言中两种最基本的表达式

1. 分号表达式：值的类型是单元类型，它实际上是一个空元组。如：

```
; -> ()  
let expr: &str = "hello";
```

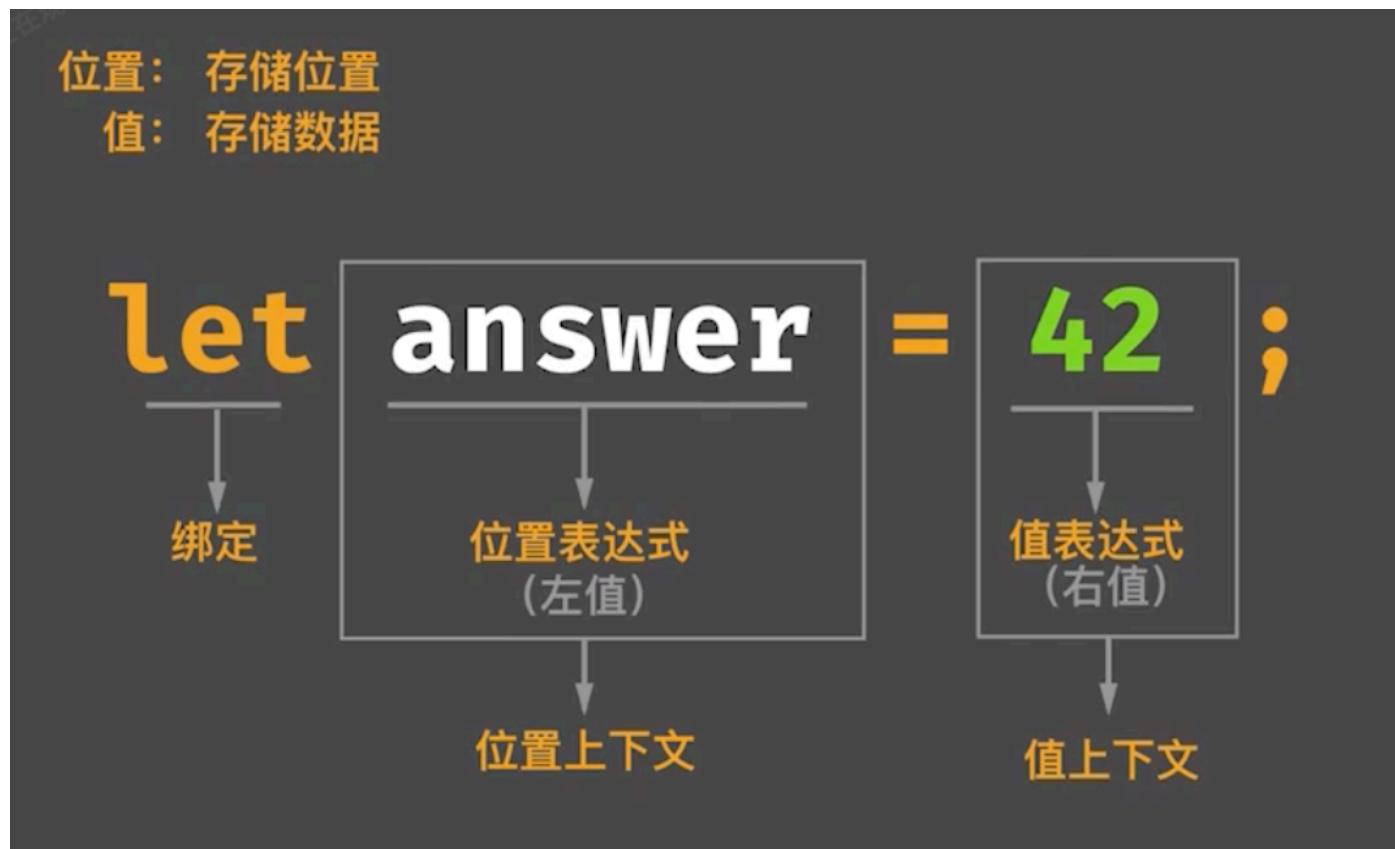
块表达式：值的类型是块中最后一个表达式的值。当块中最后一行为一个值时，块表达式的值为该值，类型是该值的类型。如：

```
let a: () = {
    let expr = "hello";
};

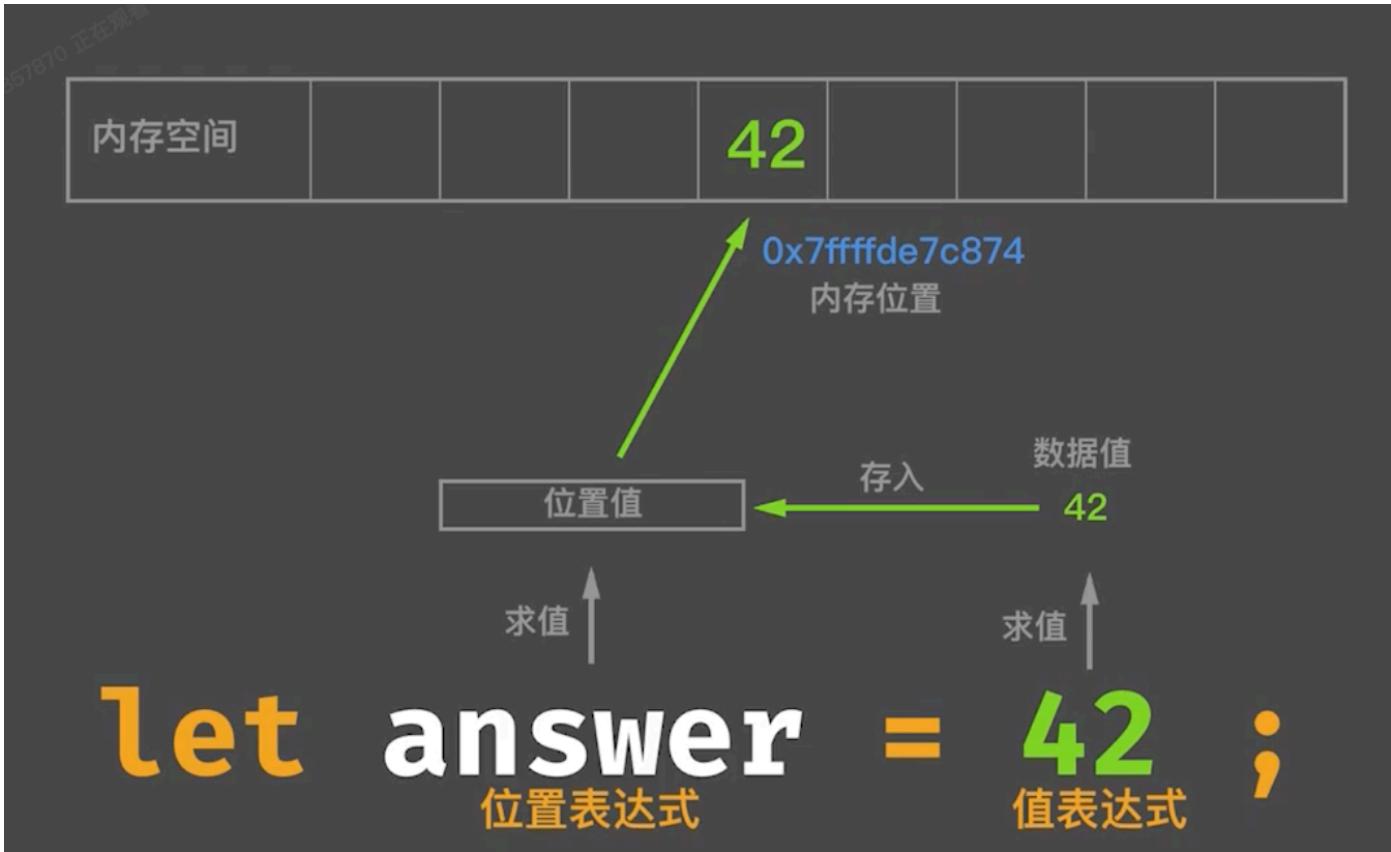
let b: &str = {
    let expr = "hello";
    expr
};
```

2.4.2 表达式分类：按内存管理

- 对于数据和变量的关系，变量处于位置区域，数据处于值区域，位置存储位置，值存储数据，二者以等号为界
- 位置表达式：位置，存储位置
- 值表达式：值，存储数据



- 表达式背后的内存管理



2.4.2.1 位置上下文

```
// 1. 赋值和复合赋值操作左侧
let mut a = 1;
a += 1;

// 2. 一元借用和解引用操作数所在区域
let a = &mut 7;
*a = 42;
// 二元借用 b:&&mut i32
let b = &a;

// 3. 字段表达式操作数所在位置
struct A {
    name: &'static str,
}
let a = A { name: "Alice" };
a.name; //位置上下文

// 4. 数组索引表达式操作数所在区域
```

```
let mut arr = [1, 2, 3];
let b = &mut arr;
arr[1];

// 5.任意隐式借用操作数所在区域
let mut v = vec![1, 2, 3];
v.push(4);

// 6.let 初始化
let a: i32;

// 7.if let/while let/match 的匹配表达式所在的区域
let dish = ("ham", "eggs");
if let ("bacon", b) = dish {
    // ("bacon",b) 就是位置上下文
    println!("bacon is served {}", b);
} else {
    println!("No bacon will be served")
}

//match (位置上下文) / while let (位置上下文) 同理

// 结构体更新语法中的base表达式
struct Point3d {
    x: i32,
    y: i32,
    z: i32,
}

let mut base = Point3d { x: 1, y: 2, z: 3 };
let y_ref = &mut base.y;

Point3d {
    y: 0,
    z: 10,
    ..base
};
```

2.4.3 所有权语义在表达式上的体现

2.4.3.1 位置表达式的移动

```

let stack_a = 42;
let stack_b = stack_a; // 位置表达式到值上下文中, 发生了copy

stack_a;

let heap_a = "hello".to_string();
let heap_b = heap_a; // 位置表达式到值上下文中, 发生了move

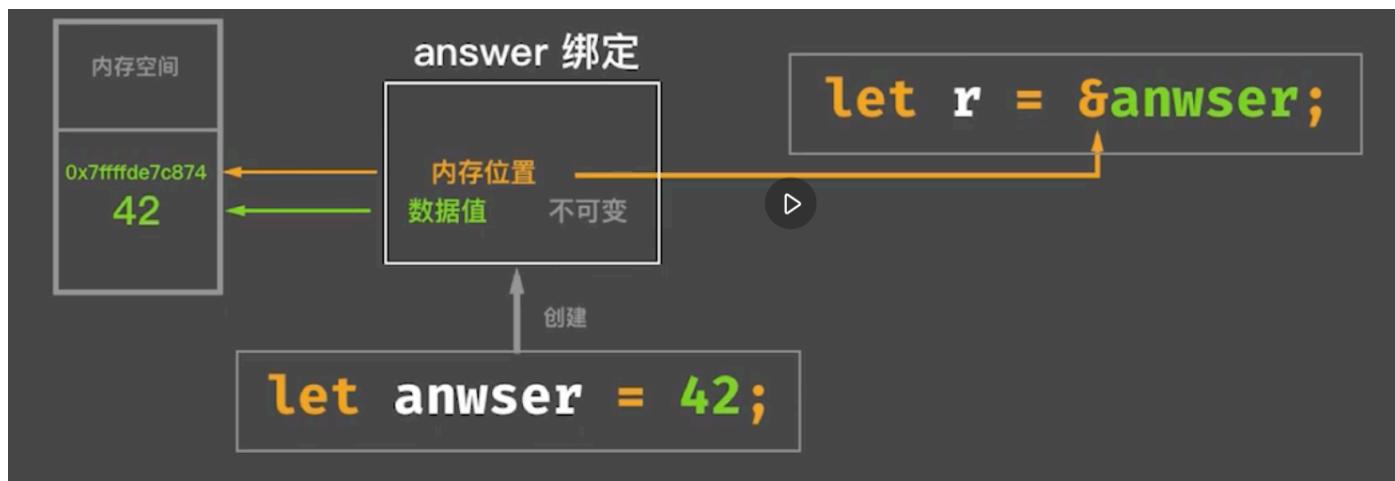
//      heap_a; error

```

2.4.4 不可变与可变

由于所有权机制, 一个内存地址只能有一个绑定

1. 不可变绑定与可变绑定
2. 不可变引用 (共享引用) 与可变引用 (独占引用)



3. Rust中与C语言一样的*mut T和 *const T 只能在Unsafe Rust中使用, Rust中的原生指针没有所有权语义

2.5 编译期计算

编译期计算 (CTFE) : 编译期函数求值, 最先由Lisp/Cpp语言支持

2.5.1 Rus编译期计算方式

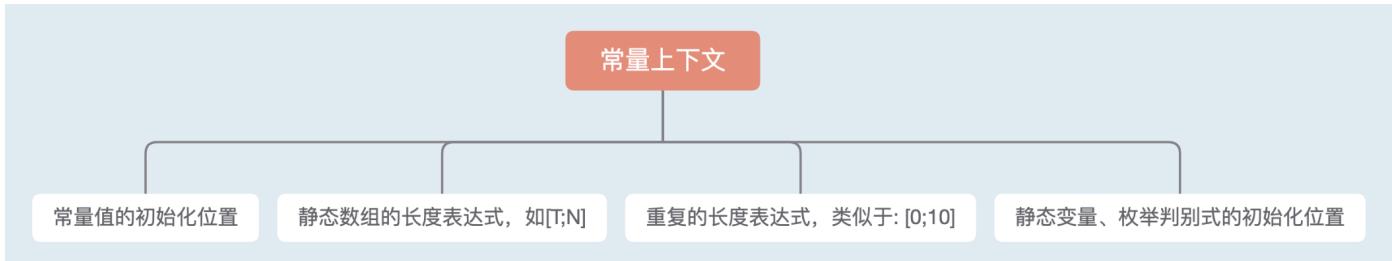
过程宏 + Build脚本 (build.rs) : 类型计算、生成代码等, 但是无法在宏代码和普通代码之间共享代码

类似Cpp中的constexpr的CTFE功能: 分为两类: 常量函数和常量泛型

2.5.1.1 常量表达式和常量上下文

在编译期对常量表达式进行求值

```
const AN: i32 = 1000; //常量表达式
```



1. 常量上下文是编译器唯一进行常量求值的地方
2. 编译期计算默认是对开发者透明的，但是了解这部分的知识能够让你对底层更有sense
3. 与常量计算相对应地一个知识点：常量传播，它是编译器的一种优化，防止运行时重复计算

2.5.1.2 常量安全

1. 理论上，Rust中的大部分表达式都可以用作常量表达式，但目前只支持常量函数，元组结构体，元组的值
2. 并不是所有常量表达式都可以用在常量上下文：比如某个数组的长度依赖于磁盘中文件内容的长度。因为编译期求值必须得到确定结果，当文件发生变化时就无法保证确定性
3. 因此rust引入了常量函数解决常量安全问题

```
// 1. 常量函数
const fn gcd(a: u32, b: u32) -> u32 {
    match (a, b) {
        (x, 0) | (0, x) => x, //
        (x, y) if x % 2 == 0 && y % 2 == 0 => 2 * gcd(x / 2, y / 2),
        (x, y) | (y, x) if x % 2 == 0 => gcd(x / 2, y / 2),
        (x, y) if x < y => gcd((y - x) / 2, x),
        (x, y) => gcd((x - y) / 2, y),
    }
}

const GCD: u32 = gcd(21, 7);

println!("{:?}", GCD);
```

2.5.1.3 编译期计算如何实现

Rust编译器内置了MIR解释器：Miri，它会执行中级中间语言中const上下文中的const代码，从而实现编译期计算
一个小知识点：无限循环用loop而不是while true

2.5.1.4 常量泛型

Rust中静态数组是二等公民，长度不同类型不同，我们无法使用统一的命名所有数组

```
// 可以定义泛型结构体
pub struct S<T, N> {
    x: T,
    y: N,
}

// 但是无法定义泛型数组
// let arr = [T; N]; // 不支持
```

问题：如何定义一个泛型静态数组，等到真正填充数据的时候，再决定数组中元素的类型以及长度？

问题的核心是：在未初始化数据结构的情况下在分配内存空间

解决方案：

1. 使用泛型结构体声明泛型参数 `T` 和常量泛型 `N`
2. 使用核心库中的联合体 `MaybeUninit` 包裹泛型参数占位
3. 用于给泛型生成一个未初始化的示例，并再构建一个泛型结构体，泛型参数分别是类型 `T` 和常量泛型。
`MaybeUninit` 用来占位

解决方案的核心是：先分配内存空间，再初始化数据结构

```
#![feature(min_const_generics)]
use core::mem::MaybeUninit;

#[derive(Debug)]
pub struct ArrayVec<T, const N: usize> {
    items: [MaybeUninit<T>; N], // 先分配内存
    length: usize,
}

fn main() {
    println!();

    let av = ArrayVec {
        items: [MaybeUninit::uninit(); 3], // 再声明数据结构
        length: 10,
    };

    println!("{:?}", av)
}
```

```
// 打印结果:  
ArrayVec {  
    items: [  
        core::mem::maybe_uninit::MaybeUninit<u32>,  
        core::mem::maybe_uninit::MaybeUninit<u32>,  
        core::mem::maybe_uninit::MaybeUninit<u32>,  
    ],  
    length: 10,  
}
```

常量泛型目前只支持

1. 一个简单的常量泛型参数，比如 `const N:usize`
2. 可以在不依赖任何类型或常量参数的常量上下文中使用表达式

保留的问题：什么时候使用常量泛型呢

```
// array_chunks 方法是基于常量泛型对数组进行分割处理  
  
let data = [1, 2, 3, 4, 5, 6];  
let sum1 = data.array_chunks().map(|&[x, y]| x * y).sum::<i32>();  
let sum2 = data.array_chunks().map(|&[x, y, z]| x * y * z).sum::<i32>();  
assert_eq!(sum1, (1 * 2) + (3 * 4) + (5 * 6));  
assert_eq!(sum2, (1 * 2 * 3) + (4 * 5 * 6));  
  
println!("{} , {}", sum1, sum2);
```

2.6 Rust 类型系统

2.6.1 类型系统目标

1. 保证内存安全
2. 保证一致性
3. 表达明确的语义
4. 零成本抽象表达能力

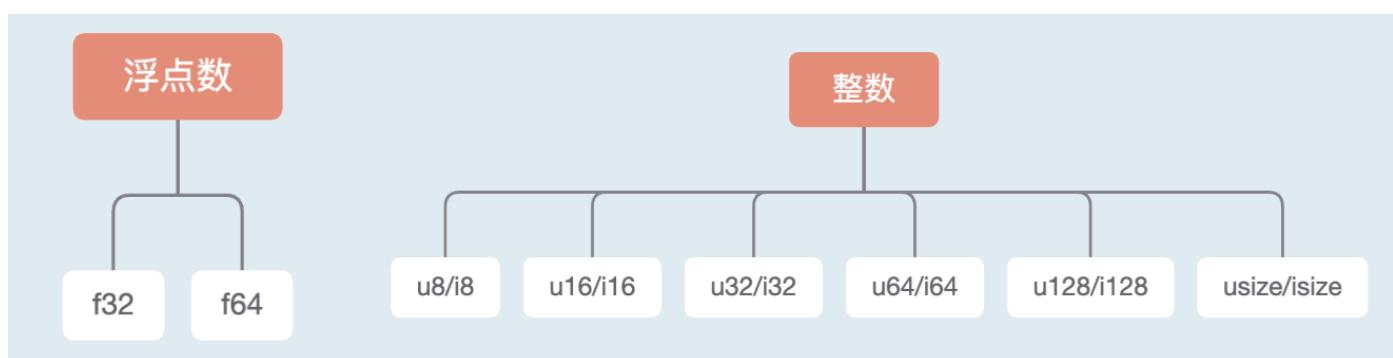
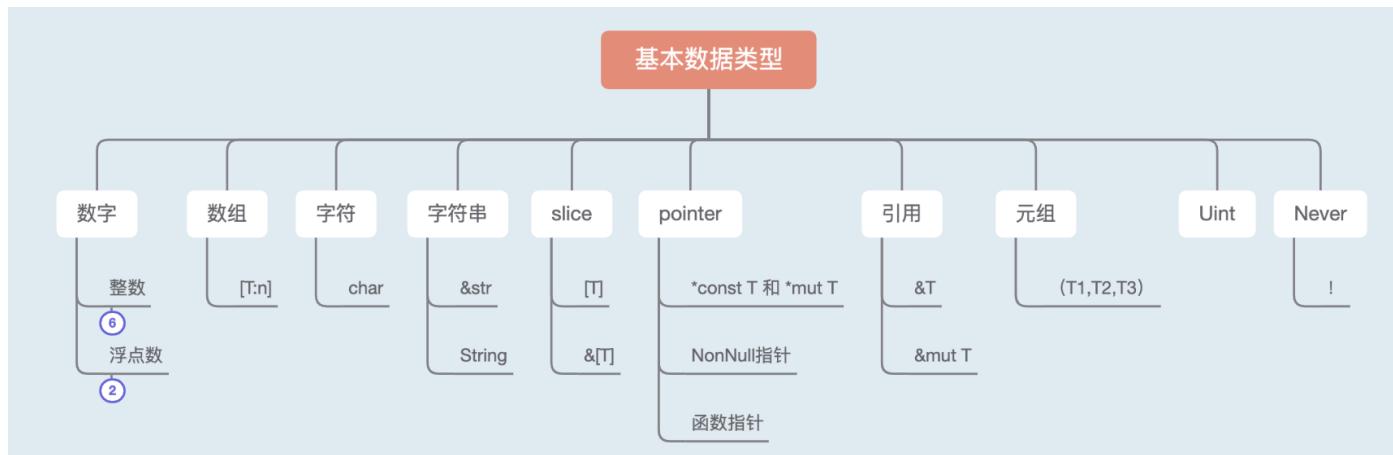
2.6.2 Rust如何实现目标

类型：在rust中，一切皆类型

行为：Rust使用trait规范了类型的行为

2.6.3 Rust数据类型

2.6.3.1 基本数据类型



特别说明

1. usize和 isize有符号和无符号指针大小类型，指针一般和计算机字长相等，32位处理器：4字节，64位处理器：8字节
2. 布尔值可以转数字，但是反过来不可以
3. 数组在Rust中是二等公民，长度不同，类型不同。等常量泛型稳定后可以晋升统一的[T;N]! 类型
4. 字符，rust中的char是unicode标量，占四个字节,对应于Rust中的u32类型。并且可以方便的转换为utf8编码的字节序列。字节序列的每一个元素是1个字节。注意对应u32类型，并不代表所占字节是4字节，所占字节仍然遵从unicode规则
5. 补充知识：

5.1 ASCII 字符集（英文字符集）：使用一个字节存储，共计128个字符

5.2 GBK:汉字字符集，占两个字节，共计两万多个，编码第一位是1

5.3 Unicode字符集：编码方案：uft-8，可变长编码方案，共分为四个长度区：1字节，2字节，3字节，4字节

汉字：1110xxxx 10xxxxxx 10xxxxxx，英文数字一个字节

```
let tao = '道';

let tao_u32 = tao as u32;
println!("{}: {}", tao, tao_u32); // 字符'道'对应的u32的值
println!("U+{:x}: {}", tao_u32, tao); // 道的Unicode 字符编码
```

```
println!("{}", tao.escape_unicode()); // 道转译后的Unicode 码点

let a = char::from(65);
println!("{}", a);

//转换16进制的码点值
if let Some(c1) = std::char::from_u32(0x9053) {
    println!("{}", c1)
}
if let Some(c2) = std::char::from_u32(36947) {
    println!("{}", c2)
}

// 并不是所有的数字都是有效的Unicode标量值
if let Some(c3) = std::char::from_u32(129010101) {
    println!("{}", c3)
} else {
    println!("invalid code")
}

use std::str;
// 将utf-8序列转换为字符串
let tao = str::from_utf8(&[0xE9u8, 0x81u8, 0x93u8]).unwrap();
println!("tao:{}", tao);

// 通过16进制码位转换为字符串
let tao = String::from("\u{9053}");
println!("{}", tao);
let unicode_x = 0x9053;
let utf_x_hex = 0xe98193;
let utf_x_bin = 0b11101001100000011001011;

println!("unicode_x: {:?}", unicode_x);
println!("utf_x_hex: {:?}", utf_x_hex);
println!("utf_x_bin: {:?}", utf_x_bin);

// 特殊字符
// 码位可能不同,但字节大小一样
// 长度可能不同,但值的大小一样

let e = 'é'; // 和 let e = 'é'; 不一样, 前者是两个unicode 码点, 后者是1个
            // let e = 'é';
let f = 'e';

let g = "é";
let h = "e";

println!("e utf-8 bytes: {}", e.len_utf8()); // 占2个字节
println!("f utf-8 bytes: {}", f.len_utf8()); // 占1个字节
```

```

println!("e value size: {}", std::mem::size_of_val(&e)); // 4字节
println!("f value size: {}", std::mem::size_of_val(&f)); // 4字节

println!("g utf-8 bytes: {}", g.len()); // 2字节
println!("h utf-8 bytes: {}", h.len()); // 1字节

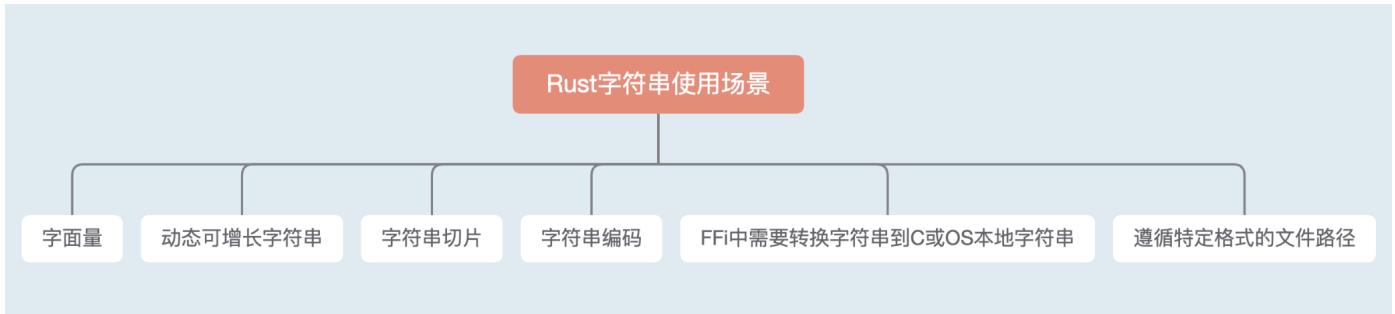
println!("g value size: {}", std::mem::size_of_val(&g)); // 16字节
println!("h value size: {}", std::mem::size_of_val(&g)); // 16字节

// emoji 只能是字符串
let s = String::from("love: ❤");
println!("emoji {}", s)

```

实现的 trait 有 Copy、Clone 等

6. 字符串，rust 中的字符串有非常多的类型，从根本上讲是为了适应不同的场景，如下：



在 Rust 中，字符串比较复杂，涉及底层内存管理知识

```

// 类型是 &str, 字符串切片的引用, 胖指针(指针和数据长度), 原属数据存放在静态存储区
let s_static_memory = "hello";

// 不可以使用未知大小的静态存储区的原始字符串
// let s = *s_static_memory;

// 类型是 String, 字符串的引用, 智能指针(指针、容量和数据长度), 原属数据存放在堆上
let s_heap_memory = String::from("hello");

// 不可以使用未知大小的堆上原始字符串
// let s = *s_heap_memory;

```

Rust 中每一个字符串都是一个 UTF-8 字节序列，实际上是一个“Vec”动态数组

6.1 两种常见类型：str（字符串切片）和String

Rust 中没有内含正则引擎，日常字符串操作通过它本身的一些方法来完成字符、字节和字符串之间的转换。还有一些定位、搜索、匹配、去除空白等方法。可以在多线程中安全的使用

6.2 String 为什么有容量，因为它基于数组

Pattern 相关的 trait 提供了同名函数不同参数的功能，可以重点看看

6.3 其他类型：

1. Cstr/Cstring 与C语言打交道
2. OsStr/OsString 与操作系统打交道
3. Path/PathBuf 与路径打交道

标准库导读三原则

1. 类型自身介绍
2. 类型自身实现的方法
3. 类型实现的trait

7 指针类型

两种原始指针：`*mut T` 和 `*const T`

NonNull指针：替代`*mut T`, 非空，并遵循生命周期类型协变规则

函数指针：指向代码而非数据，可以用于直接调用函数

8 引用与指针之别

1. 引用不为空
2. 拥有生命周期
3. 受借用检查器保护，不会发生悬垂指针等问题

8 元组

Rust中唯一的异构序列

长度不同类型不同

单元类型的唯一实例等价于空元组

当元组只有一个元素时需要在元素后加逗号，以做区分

```
// 类型是元组: (i32,)  
let a = (42,);  
// 类型是 i32  
let b = (42);
```

9 Never类型

代表不可能返回值的计算类型

在类型理论中叫底类型，不包含任何值，但是可以合一到任何其他类型。用! 表示（目前还未稳定）

2.6.3.2 自定义复合类型

2.6.3.2.1 结构体

2.6.3.2.1.1 结构体种类

```
// 1.具名结构体
struct Point {
    x:f32,
    y:f32
}

// 2.元组结构体,常用于包装基本数据类型以扩展功能
struct Pair(i32,i32);
// 当元组结构体只包含一个类型是, 称为NewType模式
// 如下对u32进行包装, 表示分数
struct Score(u32);

impl Score {
    fn pass(&self) -> bool {
        self.0 >= 60
    }
}

let s = Score(59);
assert_eq!(s.pass(), false);

// 3.单元结构体,实例就是它自身, 0大小
struct Uint;

let point = Point { x: 3.0, y: 4.0 };
let pair = Pair(1, 1);
let uint = Uint;

assert_eq!(point.x, 3.0);
assert_eq!(pair.0, 1);
```

2.6.3.2.1.2 内存对齐方式

```
// 推断结构体占12字节
// #[repr(C)] //使用属性不让编译器自动优化布局
struct A {
    a: u8,    // 占1字节,按照4字节对齐, 补3
    b: u32,   // 占4字节, 补0
    c: u16,   //占2字节, 补2
}

// 实际优化,字段重排
struct B {
    b: u32, // 计算机按照字节寻址, 指令是字节的整数倍
```

```

c: u16,
d: u8,
}

println!("{:?}", std::mem::size_of::<A>());
println!("{:?}", std::mem::size_of::<B>());

```

2.6.3.3 容器类型



2.6.3.3.1 共享容器

内部可变性：本质是把原始指针*mut 给开发者

1. 与继承式可变相对应（继承式可变，前面声明了一个不可变，紧接着又声明了可变）
2. 由核心原语UnsafeCell提供支持，UnsafeCell是Rust中唯一可以把不可变引用转为可变指针的方法
3. 基于UnsafeCell,提供了Cell和RefCell

```

### 容器Cell、RefCell、UnsafeCell
### 1. 容器Cell: 通过移进移出值来实现内部可变性
```
use std::cell::Cell;
struct Foo {
 x: u32,
 y: Cell<u32>, // 包裹实现了copy trait的类型
 z: Cell<String>, // 包裹未实现copy trait的类型
}

// 初始化一个不可变实例
let foo = Foo {
 x: 1,
 y: Cell::new(3),
 z: Cell::new("hello".to_string()),
};

assert_eq!(1, foo.x);
assert_eq!(3, foo.y.get());
// 没有实现copy的类型无法使用get方法获取内部值,可以看到Cell容器是通过移进移出值来实现内部可变性的
// assert_eq!("hello".to_string(), foo.z.get());

// 改变不可变实例
foo.y.set(100);
println!("y: {:?}", foo.y.get());
foo.z.set("world".to_string());

```

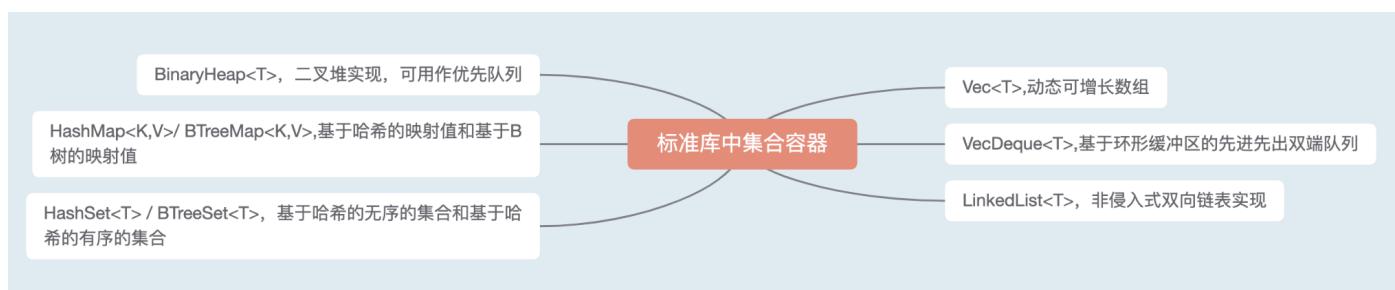
```

// 未实现copy的类型不可以使用get获取,但是可以使用into_inner获取
println!("z: {:?}", foo.z.into_inner());
// 实现了copy的类型既可以使用get获取,也可以使用into_inner获取
println!("y: {:?}", foo.y.into_inner());
```
#### 2. 容器RefCell: 通过borrow_mut实现可变性
// 主要是应用于一些未实现copy trait类型, 通过borrow获取值, 有运行时开销
```
use std::cell::RefCell;
// 使用vec! 宏创建不可变的动态可增长数组
let vec = vec![1, 2, 3, 4];
// vec.push(5); // 不能往不可变的数组中增加元素

let ref_vec = RefCell::new(vec); // 包裹变长数组
println!("{:?}", ref_vec.borrow()); // 不可变借用打印
ref_vec.borrow_mut().push(5); // 可变借用改变
println!("{:?}", ref_vec.borrow()); // 不可变借用打印
```
#### 3. 容器UnsafeCell 是上述两种容器的底层实现

```

2.6.3.3.2 集合容器



1. Vec 标准库导读

自身的方法: 转换、排序、二分搜索、组合链接 (join) 、交换、追加等,

实现的trait: Default

```

pub trait Default: Sized {
    fn default() -> Self;
}

```

Rust的内存分配器可以自定义; Vec内部是一个结构体, 还介绍了容量和重新分配的概念。按照预分配的成倍增加。不会自动缩减。存放于堆上。如果相对存放的位置进行优化, 可以使用rust- smallvec库

集合容器: collection。什么时候用哪些? 性能, 迭代器, 容量管理 (手工使用缩减) 、entry模式 (连续插入)

2. LinkedList标准库导读

增删改查, node存储数据非侵入式, 侵入式的不存储数据。建议尽量使用动态数组和双端队列

3. HashMap标准库导读

基于二次探查和SIMD查找, 数据级的并行, 就是单指令多数据查找

一般对哈希表的要求，哈希值如何产生，如何避免哈希冲突。Rust哈希算法默认是siphash，可以实现Hasher trait 替换哈希算法，如FnvHasher， 默认可以抵抗HashDos攻击。如何解决哈希碰撞，现在是Google的SwissTable实现，和C++持平。以前用的是Robinhood，但他们都基于二次探查

枚举在rust中相当于一个接口

方法：和动态数组差不多，实现trait: Extend，没有实现Drop，因为内部使用了算法hashbrown，实现了drop，涉及数据并行。还需要关注一个设计模式，entry，entry返回一个枚举（占位和空缺两种状态），非常聪明

Rust集合容器为什么没有统一的接口（trait）：缺乏功能泛型关联类型GAT的支持

2.6.3.4 泛型

在Rust中,泛型是零成本的，因为在编译期就单态化（在实际调用的位置生成具体类型相关的代码），也叫静态分发

```
fn foo<T>(x: T) -> T {
    x
}

fn main() {
    assert_eq!(foo(1), 1);
    assert_eq!(foo("hello"), "hello");

    // 上述的函数会单态化为两个不同参数类型的函数
    fn foo_1(x: i32) -> i32 {
        x
    }
    fn foo_2(x: &'static str) -> &'static str {
        x
    }

    foo_1(1);
    foo_2("2");

    // Rust根据上下文有一定的推断能力，但是推断不出来时需要手工通过turbofish指定
    // foo(1) 等价于 foo::<i32>(1);
    // foo("hello") 等价于 foo::<&'static str>("hello");
}
```

2.6.3.5 特定类型

特定类型是指专门有特殊用途的类型，Rust中有两种

1. PhantomData, 幻影类型：一般用于Unsafe rust的安全抽象或者占位。

`PhantomData<T>`不包含任何实际数据，只用来记录类型信息

通常在 Rust 的一些底层编程中，我们需要知道一个数据结构中存储的类型，但不需要存储该类型的实际数据。这种情况下，可以使用`PhantomData<T>`。

使用`PhantomData<T>`可以帮助我们保持某些类型的类型安全，同时不需要存储任何实际数据。这有助于减少内存使用，因为我们不需要分配任何内存来存储类型相关的数据。

总的来说，`PhantomData<T>`是一种辅助工具，帮助我们在不存储实际数据的情况下，知道某些数据结构中存储的数据的类型。

```
use std::marker::PhantomData;

#[derive(Debug)]
struct Container<T, U> {
    data: T,
    marker: PhantomData<U>,
}

impl<T, U> Container<T, U> {
    fn new(data: T) -> Container<T, U> {
        Container {
            data,
            marker: PhantomData,
        }
    }
}

fn main() {
    // 我们知道结构体的第二个字段是u32，但是它的值是多少我们并不在意
    let _container: Container<i32, u32> = Container::new(42);
    println!("{:?}", _container)
}
```

2. Pin, 固定类型：为了支持异步开发特意引进，防止被引用的值发生移动的类型

`Pin<T>`是 Rust 编程语言中的一种数据类型，它是一种指针类型，用于表示不可移动的指针。

在 Rust 中，通常情况下，对象的地址可能随着时间的推移而发生变化。但是，有时我们希望一个对象的地址保持不变，这时就可以使用`Pin<T>`类型。

使用`Pin<T>`类型可以防止 Rust 自动重新分配内存并改变指针的指向。这样，可以在编写需要一个不可移动的指针的代码时，更方便地管理内存。

总的来说，`Pin<T>`是一种特殊的指针类型，用于保证一个对象的地址不变，从而在某些场景下更方便地管理内存。

```
use std::pin::Pin;
```

```

use std::mem::MaybeUninit;

struct MyStruct {
    data: u32,
}

fn main() {
    let mut x = MaybeUninit::::uninit();
    let x = unsafe { x.as_mut_ptr() };
    let x = unsafe { Pin::new_unchecked(x) };
    x.as_ref().write(MyStruct { data: 42 });
    println!("{}: {}", x.data);
}

```

使用 `MaybeUninit` 类型来创建一个未初始化的内存空间，然后将其转换为指针。最后，使用 `Pin::new_unchecked` 将该指针包装在 `Pin<T>` 类型中。通过使用 `Pin<T>`，可以保证该指针指向的内存空间不会发生变化，从而避免了内存安全问题。

在 Rust 中，指针指向的内存空间通常在以下情况下会发生变化：

重新分配内存：当对象的大小或类型发生变化时，Rust 可能会重新分配内存以存储该对象。这会导致原来指向该对象的指针指向不同的内存空间。

移动：在 Rust 中，对象通常是不可移动的，因此其地址不变。但是，当移动该对象时，其指针的指向将发生变化。

释放内存：当没有对象再引用一个对象时，Rust 会释放该对象占用的内存。这样，指向该对象的指针将不再指向有效的内存空间

2.7 类型的行为

2.7.1 trait

1. trait 含义

本质上是定义了公共的方法，以便达到某个目的。任何类型想要达到某个目的，有两种方式，一种是自己定义方法去实现，另一种就是接入到 trait 系统中来，实现 trait 中一定定义好签名的方法。第二种会让代码更清楚明了和有约束性

2. trait 实现

trait 中也可以定义默认实现和定义关联类型（一般是返回值类型中的错误类型）

```

// 单个类型的解析
let four: u32 = "4".parse().unwrap();
println!("{}: {}", four);

// 元组结构体的解析
// 解析思路是先拿到结构体中的数字，然后使用 from_str 转化
use std::str::FromStr;

```

```

#[derive(Debug, PartialEq)]
struct Point(i32, i32);

#[derive(Debug, PartialEq, Eq)]
struct ParsePointError;

// 使用trait 提供的公共的方法来解析
// trait中有个方法是from_str,参数是字符串切片,返回值是目标类型实例
impl FromStr for Point {
    type Err = ParsePointError;
    fn from_str(s: &str) -> Result<Self, Self::Err> {
        // 实现过程因类型而异
        let (x, y) = s
            .strip_prefix('(')
            .and_then(|s| s.strip_suffix(')'))
            .and_then(|s| s.split_once(','))
            .ok_or(ParsePointError)?;
        let x_fromstr = x.parse::<i32>().map_err(|_| ParsePointError)?;
        let y_fromstr = y.parse::<i32>().map_err(|_| ParsePointError)?;

        // Ok()中包含了实例
        Ok(Point(x_fromstr, y_fromstr))
    }
}

let p = "(1,2)".parse::<Point>();
assert_eq!(p.unwrap(), Point(1, 2))

```

3. trait是一种特设多态（意思是一个接口多个实现）

Ad-hoc多态：一个接口多个实现

4. trait掌控了类型的行为逻辑

例如把一个变量赋值给另一个变量时，默认情况下时发生move语义，也就是发生所有权转移，原来的变量不再有数据的所有权

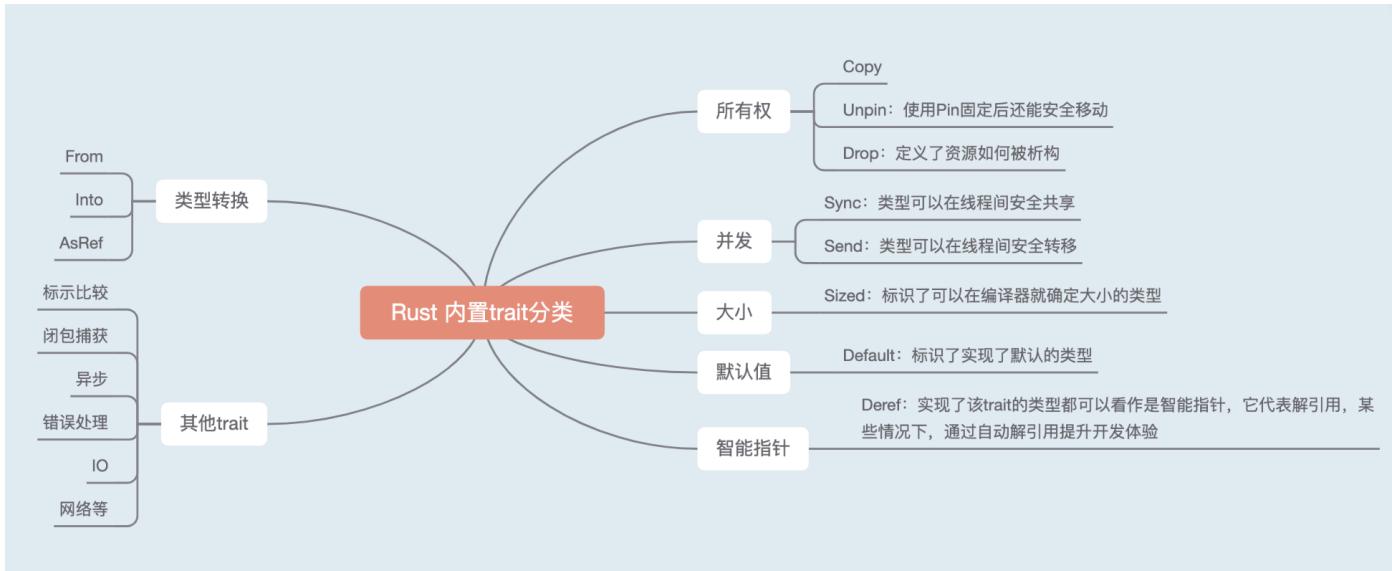
但是由于Copy trait的存在，凡是实现了Copy trait的类型，在发生上述行为时，所有权没有发生转移，而是为新的变量重新拷贝了一份数据（发生在栈上）

5. trait 理论来源

Rust类型系统遵循的是仿射类型理论，即系统中用于标识内存等资源，最多只能被使用一次。Copy trait 在整个逻辑的推理中起了很大作用

还有在rust编译器内使用了一个叫做chalk的trait系统，它是一个类似于逻辑编程语言Prolog的一个逻辑推理引擎

6. trait 分类



2.8 函数与闭包

2.8.1 函数与函数项

2.8.1.1 函数

函数的签名都是显式的

函数有三种类型：自由函数、关联函数和方法

函数自身是一种类型，值就是对应的代码

Rust语言中函数是一等公民，可以在函数间进行传递，也称高阶函数

2.8.1.2 函数项

```

struct A(i32, i32);
impl A {
    // 2. 关联函数
    fn sum(x: i32, y: i32) -> i32 {
        x + y
    }

    // 3. 方法
    fn math(&self) -> i32 {
        Self::sum(self.0, self.1) // 关联函数调用使用比目鱼符号
    }

    // 关联函数
    fn function_item(x: i32) -> i32 {
        x
    }
}

let a = A(1,2);

```

```
let x = a.math();
let y = A::sum(1, 3);

// 1.函数项构造: 类型::函数/方法名构建函数项以及自由函数的直接赋值
// 2.函数项类型: 如 fn sum(i32,i32)-> i32,就是函数签名, 同trait中的方法签名一样
let add = A::sum; //Fn item 类型
let add_math = A::math; // Fn item 类型

// 3.函数项的使用: 作为函数调用
assert_eq!(add(1, 2), A::sum(1, 2));
assert_eq!(add_math(&a), a.math());

println!("{}", x);

// 4.函数项类型本质: 0大小类型,会在类型中记录函数信息
// 好处:优化函数调用

// 5.同函数项类型一样的其他类型构造器:枚举体和单元结构体

// 5.1 函数项类型
let fn_item = A::function_item;

// 等价于
// fn function_item(_1:i32)->i32 {/* */}

// 5.1 枚举体
enum Color {
    R(i32),
    G(i32),
    B(i32),
}
// 等价于
// fn Color::R(_1: i32) -> Color {/* */}
// fn Color::G(_1: i32) -> Color {/* */}
// fn Color::B(_1: i32) -> Color {/* */}

// 5.单元结构体
struct UintStruct(i32, i32);

// 等价于
// fn UintStruct(_1: i32,_2: i32) -> UintStruct {/* */}

// 6 函数项默认实现的 trait
// Copy/Clone/Send/Sync/Fn/FnMut/FnOnce

// 7 函数项可以作为函数参数 (函数项可以当做变量) :函数项隐式转换为函数指针

// 定义一个类型别名作为返回值的类型 (RGB是三元组的类型别名)
type RGB = (i16, i16, i16);
```

```

// 自由函数
fn color(s: &str) -> RGB {
    (1, 1, 1)
}

// 参数类型是函数指针类型的自由函数
fn show(c: fn(&str) -> RGB) {
    println!("{}: {}", s, c("black"))
}

// 将函数变为函数项
let rgb = color;
// 将函数项显式转换为函数指针
let c: fn(&str) -> RGB = rgb;

// 函数指针作为另一个函数参数
show(c);
// 函数指针作为另一个函数参数, 隐式转换
show(rgb);

println!("the size of fn item {:?}", std::mem::size_of_val(&rgb)); // 0
println!("the size of fn pointer {:?}", std::mem::size_of_val(&c)); // 8

// 8 结论
// 8.1 函数项类型可以显式转换为函数指针类型, 也可以隐式转换, 但是因为携带了指针的信息, 所以要占用额外的空间
// 8.2 尽量使用函数项类型而不是函数指针, 以享受零大小类型的优化 (直接用, 不要作为参数传递)

```

2.8.1.3 函数名

函数名是一种表达式, 表达式的值是函数的相关信息, 比如类型名、参数类型名、生命周期等, 它的类型是函数项类型, 它是0大小类型

2.8.2 闭包

2.8.2.1 闭包和函数

函数只能使用传入的参数以及定义的局部变量, 无法捕获环境变量, 闭包可以

1. 闭包对环境变量的使用仍然遵循所有权机制
2. 闭包可以与函数指针互通
3. 闭包在作为函数返回值时要使用impl trait语法
4. 闭包可以捕获环境变量

```

fn counter(i: i32) -> impl FnMut(i32) -> i32 {
    // 1. 闭包与所有权
    // 闭包使用move关键字把环境变量所有权转移到闭包内
    // 具体执行copy还是move语义需要看具体的类型

```

```

let s1 = "hello".to_string();
move |s2: &str| s1 + s2;
// println!("{}:{}", s1); // 不可用, move语义

// 2. 闭包类型与函数指针类型
// 某闭包类型: |i32| -> i32, 同函数指针非常相似
// 某函数指针类型: fn(i32) -> i32

// 3. 闭包与函数指针互通 (闭包作为参数)

type RGB = (i32, i32, i32);
fn show(c: fn(&str) -> RGB) {
    println!("{}:{}", c("black"));
}

// 定义闭包: 类型 |&str| -> (i32, i32, i32), 实现了 `Fn(&str) -> RGB` trait
let c = |s: &str| (1, 2, 3);
show(c);

// 4. 闭包作为返回值
// 因为闭包是基于Trait实现的, 所以闭包作为返回值时使用的是impl trait语法
// 返回值是i32 trait的类型, 其中 FnMut(i32) -> i32 这一整块作为一个trait, 属于静态分发
// impl FnMut(i32) -> i32 代表返回的是一个实现了FnMut(i32)
let closure = move |n| n + i;
closure

}

let mut f = counter(21);
assert_eq!(42, f(21))

```

2.8.2.2 闭包实现原理

1. Rust闭包的实现与所有权机制在语义上保持了统一。闭包的三种使用场景与所有权语义三件套相匹配
2. 闭包实际上是编译器的语法糖, 也就是说, 当创建一个闭包时, 编译器会解析闭包, 并且生成一个匿名结构体, 该结构体有个泛型变量, 主要用于存储捕获的自由变量

```

// 请将下列模块属性放置在执行文件顶部
#![feature(unboxed_closures, fn_traits)]
// 按使用场景

// 1. 未捕捉环境变量 对应所有权
let c1 = || println!("hello");
c1();

// 等价于创建了一个闭包结构体, 并未闭包结构体实现了 call_once方法
// 对闭包的调用实际上是对相应trait中的方法进行调用, 但使用的名字不同, 类似在使用函数项一样
// 注意call_once方法的第一个参数是self, 代表它会消耗结构体, 需要拥有所有权

struct Closure1<T> {

```

```

    env_var: T,
}

/*
### 标准库 FnOnce trait的定义
pub trait FnOnce<Args>
where
Args:Tuple, {
    type Output;
    extern "rust-call" fn call_once(mut self, args: Args) -> Self::Output;
}
*/
// 为类型实现trait
impl<T> FnOnce<()> for Closure1<T> {
    type Output = ();
    extern "rust-call" fn call_once(self, args: ()) -> () {
        println!("hello");
    }
}

// 调用
let c1 = Closure1 { env_var: () };
c1.call_once();

// 2. 可修改环境变量 对应可变借用 &mut T
let mut arr = [1, 2, 3];
let mut c2 = |i| {
    arr[0] = i;
    println!("{}:{}", i, arr)
};

c2(100);

// 等价于
// 继承式的实现实际上是所有权一致性的体现
// 闭包实例至少需要一个消耗自身的方法

struct Closure2 {
    env_var: [i32; 3],
}

/*
### 标准库 FnOnce trait的定义
pub trait FnOnce<Args> {
    type Output;
    extern "rust-call" fn call_once(mut self, args: Args) -> Self::Output;
}

```

```

*/
// 为类型实现 FnOnce trait
impl FnOnce<(i32,)> for Closure2 {
    type Output = ();
    extern "rust-call" fn call_once(mut self, args: (i32,)) -> () {
        self.env_var[0] = args.0;
        println!("{}:?", self.env_var);
    }
}

/*
### 标准库 FnMut trait的定义
pub trait FnMut<Args>:FnOnce<Args> {
where
    Args:Tuple,
    extern "rust-call" fn call_mut(&mut self, args: Args) -> Self::Output;
}
*/
// 为类型实现 FnMut trait
impl FnMut<(i32,)> for Closure2 {
    extern "rust-call" fn call_mut(&mut self, args: (i32,)) -> () {
        self.env_var[0] = args.0;
        println!("{}:?", self.env_var);
    }
}

// 调用

let arr2 = [1, 2, 3];
let mut c2 = Closure2 { env_var: arr2 };
c2.call_mut((0,)); //可变引用调用
c2.call_once((1,)); //消耗式调用

// 3. 未修改环境变量 对应不可变借用 &T
let answer = 42;
let c3 = || {
    println!("{}:?", answer);
};

// 等价于

struct Closure3 {
    env_var: i32,
}

```

```
/*
### 标准库 FnOnce trait的定义
pub trait FnOnce<Args>
where
Args:Tuple, {
    type Output;
    extern "rust-call" fn call_once(mut self, args: Args) -> Self::Output;
}
*/



// 为类型实现 FnOnce trait
impl FnOnce<()> for Closure3 {
    type Output = ();
    extern "rust-call" fn call_once(mut self, args: ()) -> () {
        println!("{:?}", self.env_var);
    }
}

/*
### 标准库 FnMut trait的定义
pub trait FnMut<Args>:FnOnce<Args> {
where
Args:Tuple, {
    extern "rust-call" fn call_mut(&mut self, args: Args) -> Self::Output;
}
*/




// 为类型实现 FnMut trait
impl FnMut<()> for Closure3 {
    extern "rust-call" fn call_mut(&mut self, args: ()) -> () {
        println!("{:?}", self.env_var);
    }
}

/*
### 标准库 Fn trait的定义
pub trait Fn<Args>:FnMut<Args>
where
Args:Tuple, {
    extern "rust-call" fn call(&self, args: Args) -> Self::Output;
}
*/




impl Fn<()> for Closure3 {
    extern "rust-call" fn call(&self, args: ()) -> () {
        println!("{:?}", self.env_var);
    }
}
```

```
let mut c3 = Closure3 { env_var: 42 };
c3.call(()); // 不可变引用
c3.call_mut(()); // 可变引用
c3.call_once(()); // 消耗式调用
```

2.8.2.3 闭包的类型

1. 没有捕获变量，则实现 FnOnce
2. 修改捕获变量，则实现 FnMut
3. 未改捕获变量，则实现 Fn

2.8.2.4 特殊情况

1. 编译器会把 FnOnce 当成 fn(T) 函数指针区看待
2. Fn/FnMut/FnOnce 关系依次继承，对应所有权语义三件套
3. 唯一不可变借用

2.8.2.5 逃逸闭包和非逃逸闭包

```
```
// 逃逸闭包
fn c_mut() -> impl FnMut(i32) -> [i32; 3] {
 let mut arr = [1, 2, 5];
 move |n| {
 arr[2] = n;
 arr
 }
}

let i = 42;

let mut arr_closure = c_mut();
println!("{:?}", arr_closure(i));

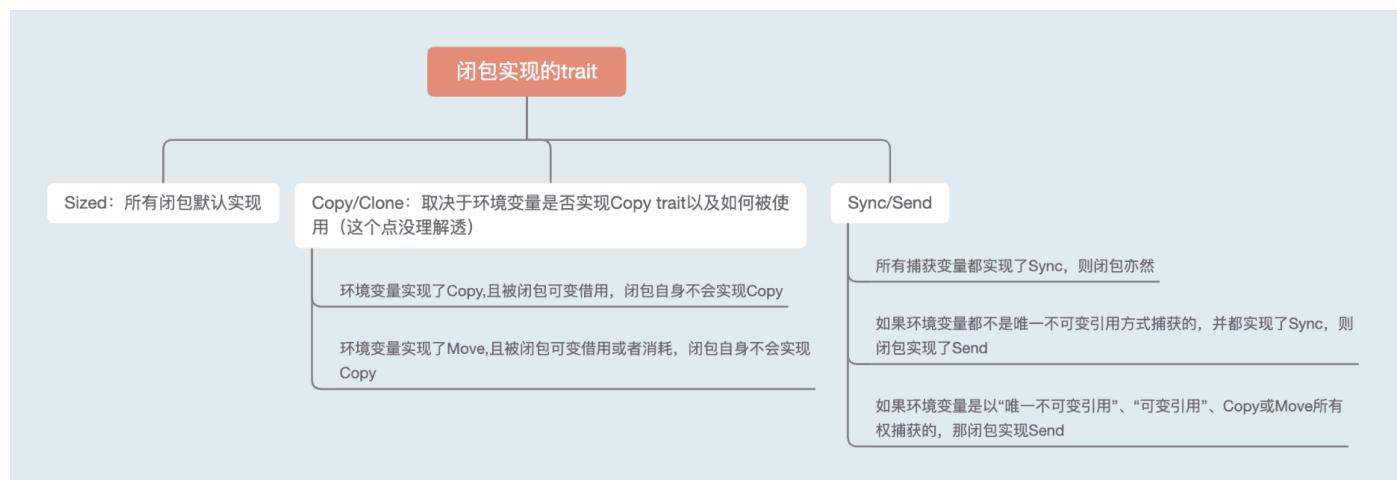
// 被捕获类型不支持 Copy，无法返回闭包，主要是为了防止悬垂引用

/*
fn c_mut2() -> impl for<'a> FnMut(&'a str) -> String {
 // 当闭包捕获了未实现 Copy trait 的类型时，无法返回
 let mut s = "hello".to_string();
 move |i| {
 s += i;
 s
 }
}
*/

```

## 2.8.2.6 闭包实现的trait

我们已知闭包会生成匿名结构体，那默认实现了哪些trait呢



只有当闭包的捕获列表中的所有变量都实现了 `clone` 和 `Copy` 时，闭包才会实现这两个 trait

```
// 闭包自身实现了Fn Copy trait
fn foo<F: Fn() + Copy>(f: F) {
 f()
}

let s = "hello".to_owned();

// 不可变借用
let f = || {
 println!("{}" , s);
};

foo(f);

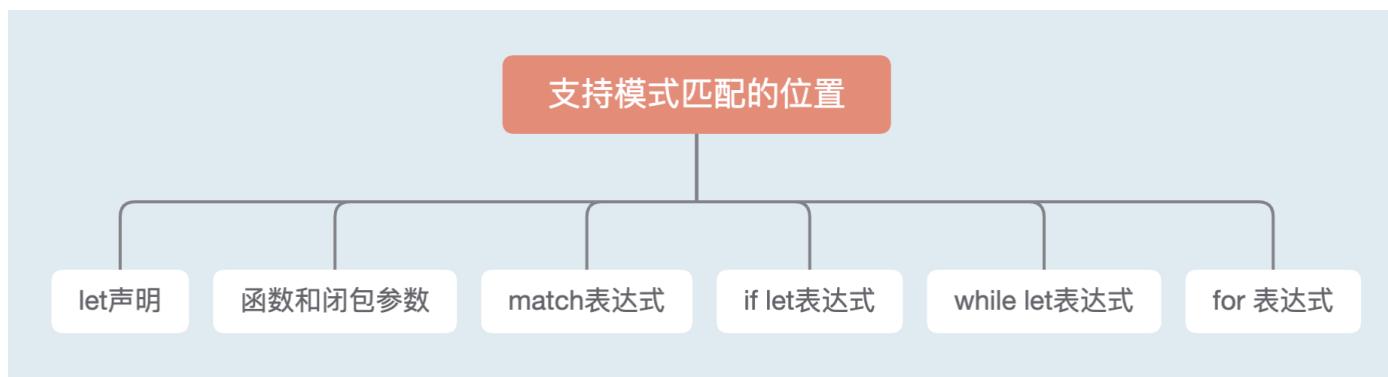
// 消耗
let g = move || {
 println!("{}" , s);
};

//foo(g); // 未实现copy trait
```

## 2.9 模式匹配

模式匹配是一种结构性的解构与构造的语义相对

## 2.9.1 模式匹配位置



## 2.9.2 模式匹配的两种类型

1. 可辩驳
2. 不可辩驳

```
// 1. let 声明中的匹配
struct Point {
 x: i32,
 y: i32,
}

let (a, b) = (1, 2);

let Point { x, y } = Point { x: 3, y: 4 };

assert_eq!(1, a);
assert_eq!(2, b);
assert_eq!(3, x);
assert_eq!(4, y);

// 2. 函数与闭包参数

fn sum(x: String, ref y: String) -> String {
 x + y
}

let s = sum("1".to_owned(), "2".to_owned());
assert_eq!(s, "12".to_owned());

// 辅助理解 ref

{
 let a = 42;
 let ref b = a;
 let c = &a;

 assert_eq!(b, c);
```

```

let mut a = [1, 2, 3];
let ref mut b = a;

b[0] = 0;

assert_eq!(a, [0, 2, 3])
}

// 3. match 表达式

fn check_option(opt: Option<i32>) {
 match opt {
 Some(p) => println!("has value {:?}", p),
 None => println!("has no value"),
 }
}

/*
fn hand_result(res: i32) -> Result<i32, dyn Error> {
 do_something(res)?;

 // 问号等价于

 match do_something(res) {
 Ok(o) => Ok(o),
 Err(e) => return SomeError(e),
 }
}
*/

```

```

let arr = [1, 2, 3];
match arr {
 [1, ..] => "start with one",
 [a, b, c] => "not start with one",
};

let v = vec![1, 2, 3];
match v[..] {
 [a, b] => "not match",
 [a, b, c] => "matched",
 _ => "",
};

// if let 表达式

let x: &Option<i32> = &Some(3);

// 编译器自动使用ref

```

```
if let Some(y) = x {
 y;
}
```

## 2.10 智能指针

### 2.10.1 在堆上分配内存：Box

从语义上Rust的类型分为值语义和指针语义。存储在栈上的就是值语义，在语义层面上就是一种值。动态字符串和动态数组会在运行时增长，它们实际上属于指针语义，传递时传递的是存储在栈上的指针而不是全部数据

Box是Safe Rust 中唯一的堆内存分配方式

```
let x: Box<i32> = Box::new(42);
// 通过解引用来获取所包裹的值，指针都可以解引用
let y = *x;

assert_eq!(y, 42)
```

### 2.10.2 Box 内存管理机制

借鉴了Cpp的RAII, Box实现了Drop trait。当变量离开作用域时，自动调用析构函数（drop函数）销毁值

```
// 标准库中的drop实现，编译器的行为
/*
unsafe impl<#[may_dangle] T: ?Sized> Drop for Box<T> {
 fn drop(&mut self) {
 FIXME:Do nothing, drop is currently performed by compiler
 }
}
*/
```

### 2.10.3 智能指针

在Rust中，trait决定了类型的行为。所以智能指针和Deref trait、Drop trait相关

二者都实现或者实现其一都是智能指针，所以智能指针在Rust中有两种语义，自动解引用（提升开发体验）和自动管理内存（安全无忧）

只实现Deref trait：拥有指针语义，Deref赋予了类型的指针行为，通常在Rust中代表了Move语义，基本是分配在堆上的数据

只实现Drop trait：拥有内存自动管理机制，Deref赋予了类型的析构行为

### 2.10.3.1 智能指针与Deref trait

```
// 1. 自动解引用 点调用操作
// 自定义一个类型
#[derive(Copy, Clone)]
struct User {
 name: &'static str,
}

impl User {
 fn name(&self) {
 println!("{}: {:?}", self.name);
 }
}

// 调用

let u = User { name: "Alex" };
// 原来的调用方式

println!("{}: {}", u.name);
// 使用自定义的智能指针包裹
let y = MySP::new(u);

// 包裹后的调用方式
// 这里智能指针实际上自动进行了解引用，获取了里面的值，然后用值进行关联函数调用

println!("{}: {}", y.name);
// 手动解引用
let z = *y;

println!("{}: {}", z.name);

// 结论：使用类型直接调用字段 = 智能指针解引用调用 = 手动解引用调用

// 2. 自动解引用 函数参数
fn takes_str(s: &str) {
 println!("{}: {}", s);
}

let s = String::from("hello");
// String 也是一个智能指针，它包裹了 str
// 自动解引用为原始类型str后要再加&

// 调用
takes_str(&s);

// 标准库中为String类型实现了Deref trait
/*
```

```

impl ops::Deref for String {
 type Target = str;

 #[inline]
 fn deref(&self) -> &str {
 unsafe { str::from_utf8_unchecked(&self.vec) }
 }
}

// 自动解引用需要注意的地方
// 使用*x 解引用等价于 * (x.deref)

let s = Box::new("world");
let ref_s1 = *s;
let ref_s2 = *(s.deref());

assert_eq!(ref_s1, ref_s2);

// 自动解引用等价于 x.deref()

```

### 2.10.3.2 标准库中的智能指针

Rc<T>和Arc<T>(共享所有权容器, 实际内部有个引用计数器, 分别用在单线程和多线程) : Drop and Deref, 实现的目的和Box<T>相同

#### 标准库中的智能指针

Box<T>: Drop and Deref

Vec<T>和String: Drop and Deref

```

// 标准库中给泛型T实现的 Deref trait
/*
impl<T: ?Sized> const Deref for &T {
 type Target = T;

 fn deref(&self) -> &T {
 *self
 }
}

impl<T: ?Sized> !DerefMut for &T {}

** 在日常开发中非常实用
** 当我们拥有可变引用T时如果还想使用T, 则可以自动解引用, 比如点调用
impl<T: ?Sized> const Deref for &mut T {
 type Target = T;

 fn deref(&self) -> &T {

```

```
*self
}
}
*/
```

## 2.11 迭代器

### 2.11.1 迭代器trait

迭代器和Rust中的集合类型密切相关

1. 是设计模式中的一种行为模式
2. 与集合使用，在不暴露集合底层的情况下遍历集合元素
3. 将集合的遍历行为抽象为单独的迭代对象（将行为抽象为对象）

迭代器分为外部迭代器和内部迭代器，for循环实际上是外部迭代器的一个语法糖

```
// 迭代器trait
trait Iterator {
 type Item;

 fn next(&mut self) -> Option<Self::Item>;
}

// 外部迭代器语法糖for循环，相当于迭代器的next方法
// vec实现了迭代器trait
let v = vec![1, 2, 3, 4, 5];
{
 // 使用into_iter方法获得迭代器
 let mut _iterator = v.into_iter();
 loop {
 // match 匹配每一次的迭代结果
 match _iterator.next() {
 Some(i) => {
 println!("{} ", i);
 }
 None => break,
 }
 }
}

// 使用for循环遍历
let v = vec![1, 2, 3, 4, 5];
for i in 0..v.len() {
 println!("{} ", v[i]);
}

// 自定义的内部迭代器（不是主要的模式）
trait InIterator<T: Copy> {
```

```

// 指定约束是为了把闭包作为参数传递
fn each<F: Fn(T) -> T>(&mut self, f: F);
}

impl<T: Copy> InIterator<T> for Vec<T> {
 fn each<F: Fn(T) -> T>(&mut self, f: F) {
 let mut i = 0;
 while i < self.len() {
 self[i] = f(self[i]);
 i += 1;
 }

 // 等价于
 // for i in 0..self.len() {
 // self[i] = f(self[i]);
 // }
 }
}

let mut v = vec![1, 2, 3];
v.each(|i| i * 3);
assert_eq!([3, 6, 9], &v[..3])

```

## 2.11.2 标准库导读

为集合类型实现迭代器时只需要实现next方法

迭代器有三种类型: iter () &T, iter\_mut () &mut T, into\_iter () T, 对应所有权三种语义

迭代器适配器模式: 允许在迭代的时候以不同的方式迭代: 如map变迭代边映射, 还有take和filter, chain。把原来的迭代器进行封装

迭代器trait: 扩展, 消费, 两头迭代, FromIterator (和消费者配合) 等

## 2.11.3 第三方库

itertools

## 2.12 模块

1. 语法集合
2. 模块是一种软件设计思想, 降低耦合, 便于维护
3. Rust中模块用于分割代码

在rust中模块可以使用mod关键字定义, 也可默认使用单个文件作为模块

同级模块使用crate

父级模块使用super

包外模块之间使用包名

模块的可见性自定义

## 2.12.1 模块与属性

```
#[path = "foo.rs"]
mod c
// 找 c.rs

// 找inline/inner.rs
mod inline {
 #[path = "other.rs"]
 mod inner;
}

// 找路径thread_files/local_data.rs
#[path = "thread_files"]
mod thread {
 #[path = "tls.rs"]
 mod local_data
}
```

## 2.13 包管理器Cargo

```
[package]
name = "from-principle-to-practice"
version = "0.1.0"
edition = "2021"

See more keys and their definitions at https://doc.rust-
lang.org/cargo/reference/manifest.html

[dependencies]
```

1. 在rust中， package/crate都是指包， crate时编译单元
2. package包含多个crate
3. crate是实际的编译单元
4. codegen-uint: 每个crate在编译时默认被LLVM IR切割为16份， 方便并行编译

## 2.13.1 Cargo工作

没有依赖地狱问题

基本包结构



```
└── Cargo.lock
└── Cargo.toml
└── src/
 └── lib.rs
 └── main.rs
 └── bin/
 └── named-executable.rs
 └── another-executable.rs
 └── multi-file-executable/
 └── main.rs
 └── some_module.rs
└── benches/
 └── large-input.rs
 └── multi-file-bench/
 └── main.rs
 └── bench_module.rs
└── examples/
 └── simple.rs
 └── multi-file-example/
 └── main.rs
 └── ex_module.rs
└── tests/
 └── some-integration-tests.rs
 └── multi-file-test/
 └── main.rs
 └── test_module.rs
```

## — test\_modutle.rs

### 2.13.2 toml 配置文件

语义明显、无歧义的配置文件格式语言格式

```
这是一个 TOML 文档

title = "TOML 示例"

[owner]
name = "Tom Preston-Werner"
dob = 1979-05-27T07:32:00-08:00

[database]
enabled = true
ports = [8000, 8001, 8002]
data = [["delta", "phi"], [3.14]]
temp_targets = { cpu = 79.5, case = 72.0 }

[servers]

[servers.alpha]
ip = "10.0.0.1"
role = "前端"

[servers.beta]
ip = "10.0.0.2"
role = "后端"
```

### 2.13.3 Cargo 命令与工具介绍

```
cargo check 静态检查当前crate及其依赖项
cargo build 静态检查和编译
cargo run 检查+构建+执行
cargo clean 清除构建文件
cargo doc 生成文档
```

常用工具

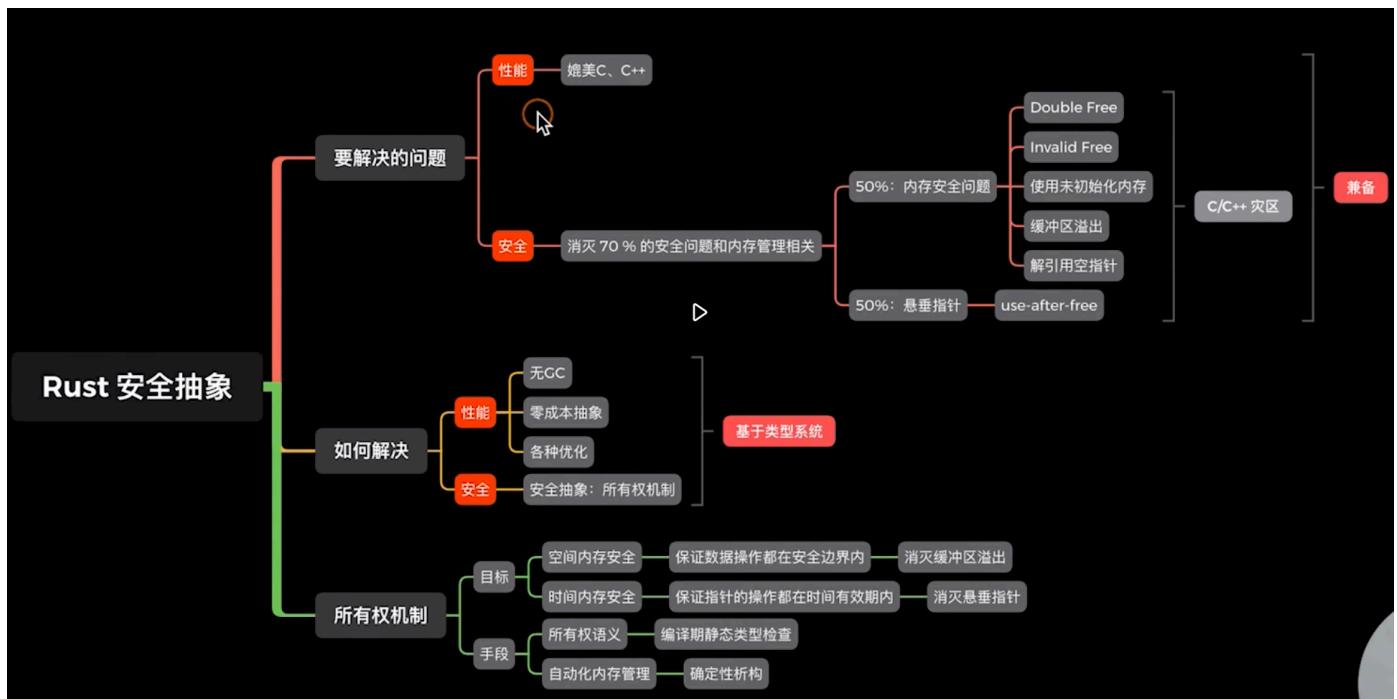
```
cargo fix 修复warning
cargo add
cargo audit 维护漏洞数据库，检查以来漏洞
cargo clippy 静态分析坏代码
cargo fmt 格式化代码
cargo expand 展开宏
```

更多内容请查看Cargo book

## 3 Rust语言核心

### 3.1 Rust语言架构

#### 1. 安全抽象和范式抽象



#### 2. 类型系统：保证程序安全

#### 3. 资源管理（内存管理）

#### 3.1.1 虚拟地址空间

虚拟内存用户地址空间

栈 stack



堆 heap

BSS

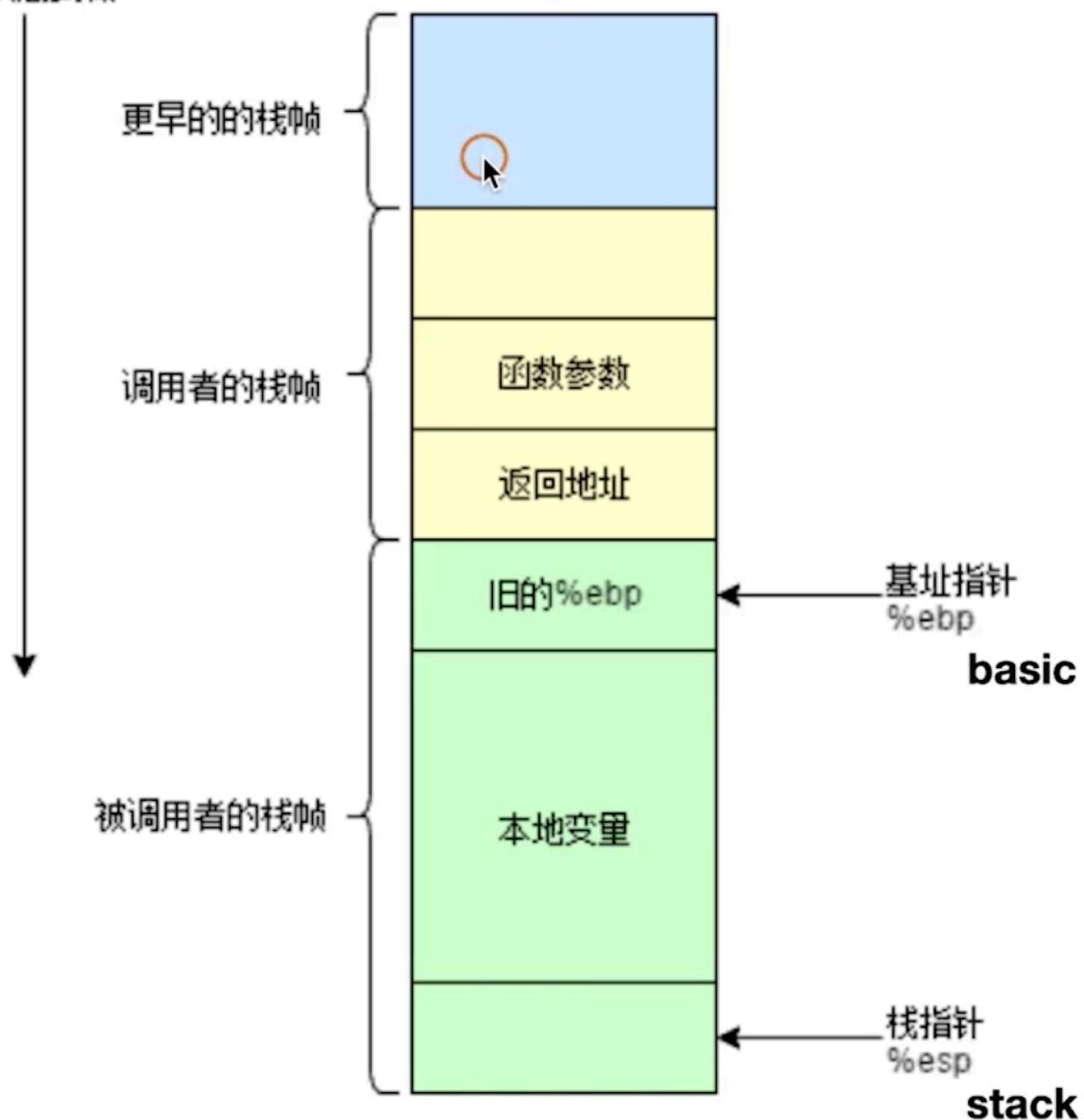
数据段

代码段

### 3.1.2 函数调用栈



地址从高到低



函数调用栈实例

```
let answer = "42";
let no_answer = answer;
println!("{:?}", answer); //可用

let answer = String::from("42");
let no_answer = answer;
```

中级中间语言

```
// MIR
```

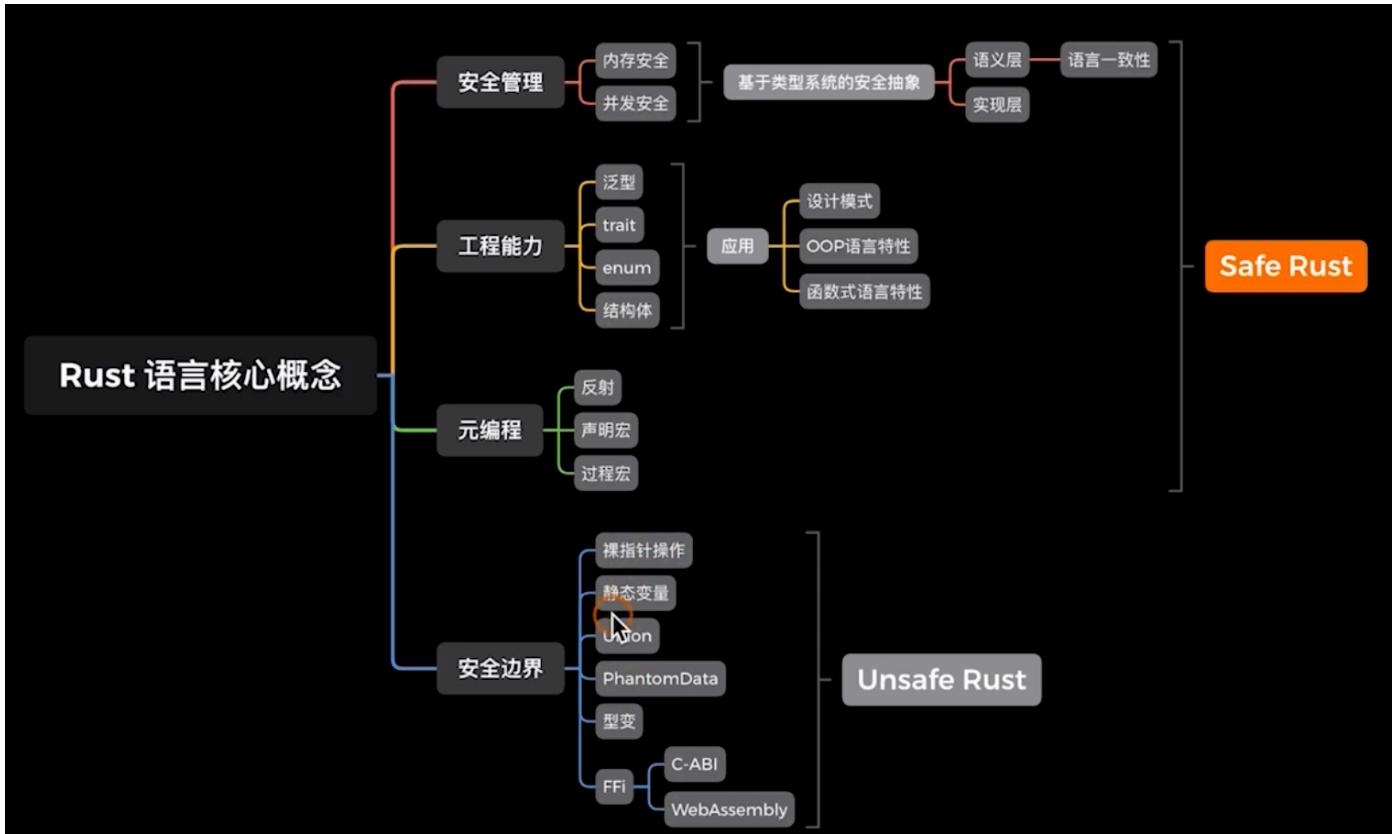
```
// 函数调用栈
// 运行结束时,最后一个会先被清除
// 先进后出
/*
let _1: &str;
scope 1 {
 debug answer => _1;
 let _2: &str;
 scope 2 {
 debug no_answer => _2;
 let _3: std::string::String;
 scope 3 {
 debug answer => _3;
 let _4: std::string::String;
 scope 4 {
 debug no_answer => _4;
 }
 }
 }
}
}
```

### 3.1.3 Rust与其它语言内存管理区别

1. C: 纯手工管理 (缺乏安全抽象模型)
2. C++: 手工管理 + 确定性析构 (缺乏安全抽象模型)
3. GC语言: 垃圾回收 (性能差)
4. Rust语言: 考虑性能, 借鉴Cpp的RAII资源管理方式, 考虑安全: 增加所有权语义

## 3.2 Rust核心概念

### 3.2.1 核心概念



### 3.2.2 要掌握的内容

1. 掌握所有权语义
2. 领略Rust的工程能力
3. 掌握元编程能力
4. 正确认识Unsafe Rust

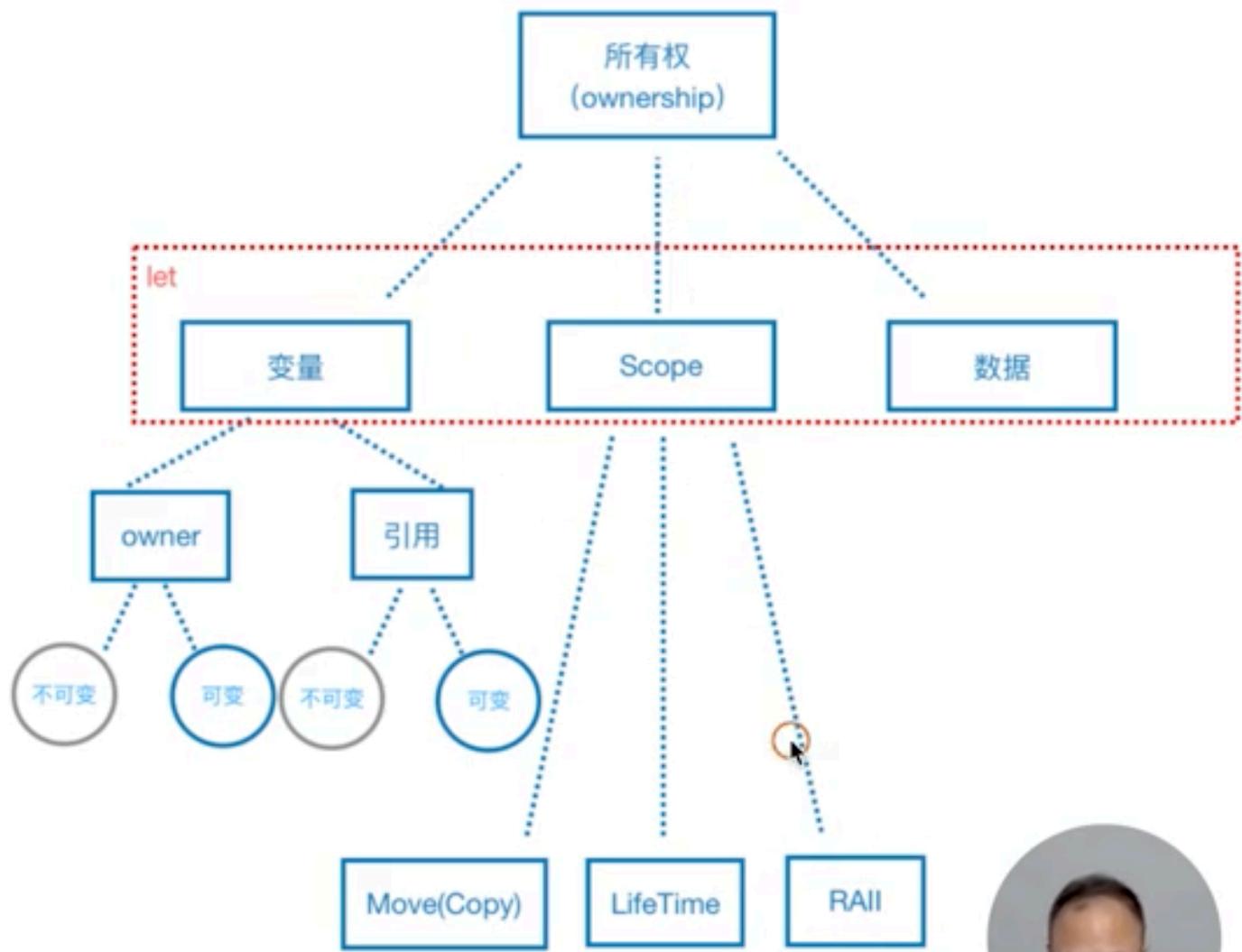
## 3.3 内存安全：所有权

### 3.3.1 语义模型

有两种: Copy和Clone

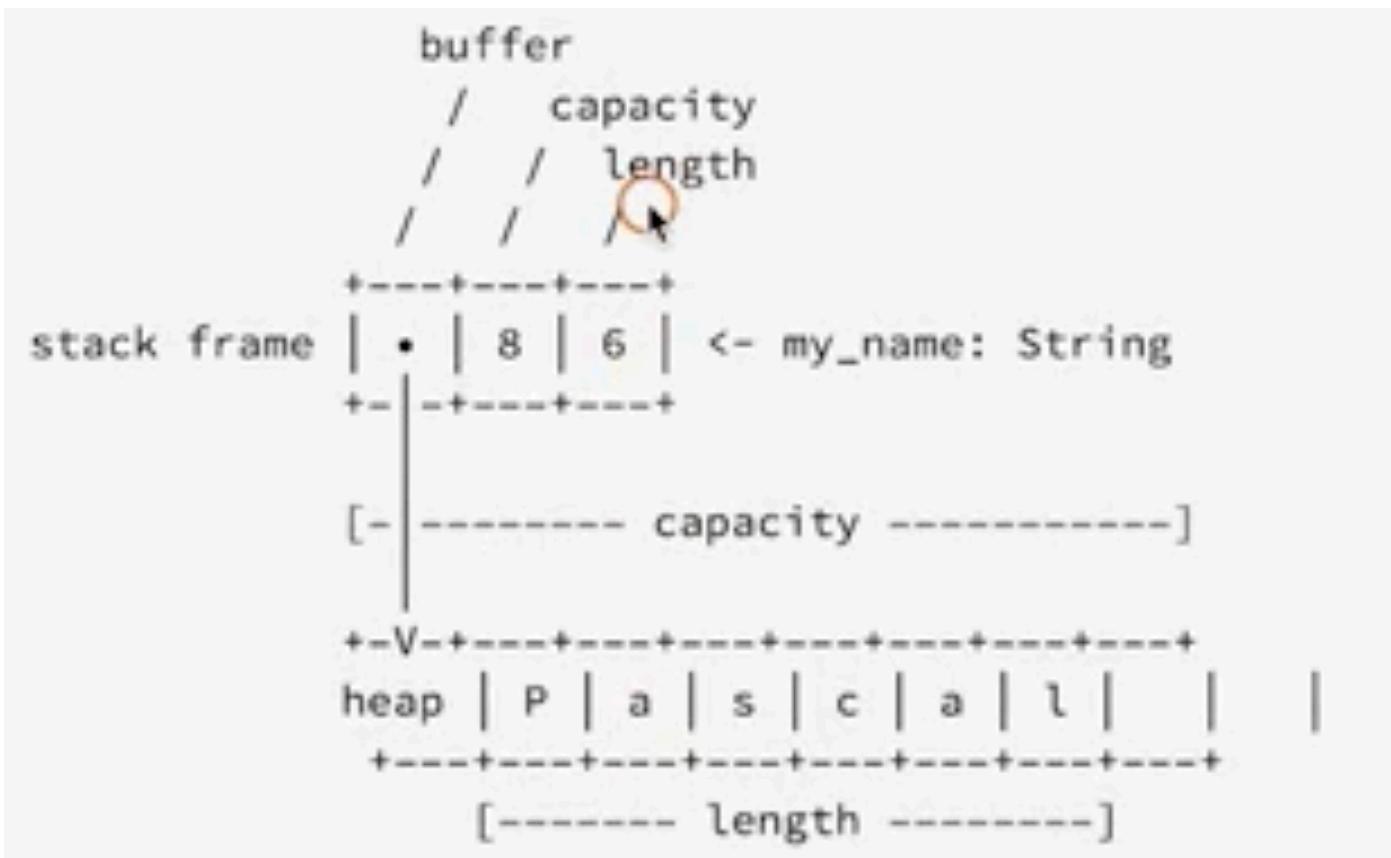
当声明一个变量时，这个变量会拥有所有权，绑定一段生命周期以及绑定一个数据，这个变量是所有权的拥有者，它可以被使用（所有权转移）或者借用（使用权转移）。当它进入到新的scope时是move或者拷贝，引用的话受原变量声明周期的约束，RALL内存管理机制通过Scope（有所有权的变量才有权利管理释放内存）管理内存

```
let answer = 42;
```

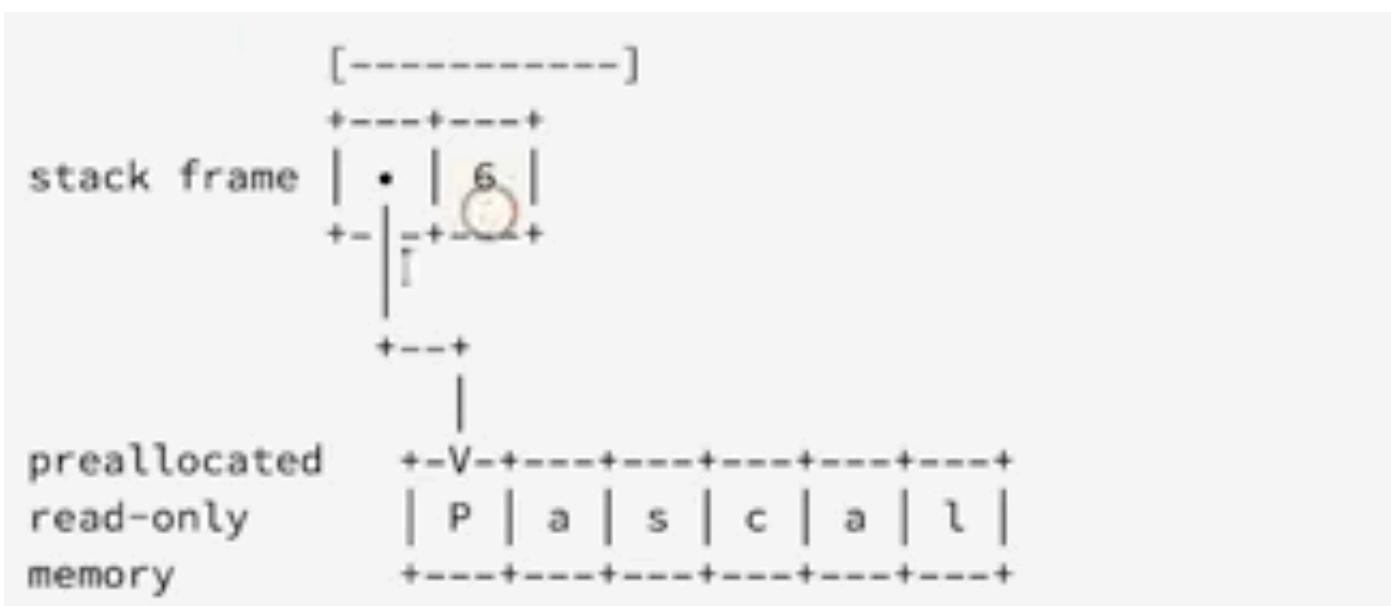


### 3.3.1.1 Copy和Copy trait

String的结构



&str的结构



基础数据类型：基本都实现了Copy trait

自定义类型：结构体不会实现Copy，需要手动通过派生宏实现，并且同时需要Clone trait；当结构体内部的成员类型没有实现copy时，结构体也不能实现Copy，枚举同理

注意：&mut T 没有实现Copy类型，&T实现了Copy

### 3.3.1.2 Move与析构

move的本质是把变量进行了未初始化标记而不是立刻丢弃

不同的情况下，变量析构的顺序可能不同，本质上是和内存安全相关的

## 3.3.2 借用检查

### 3.2.1.1 词法作用域和非词法作用域

学习词法作用域和非词法作用域借用检查

非词法作用域检查颗粒度更细，在mir层级

```
// 1 词法作用域
// 一个函数的块表达式对应一个栈帧 stack frame
// 栈帧的特点是函数调用完会自动清空
// 词法作用域对应栈帧
// 基本上词法作用域等于生命周期
let mut v = vec![];
v.push(1);

{
 // println!("{:?}", v[0]);
 v.push(2);
}
// mir 中每一个scope都代表一个词法作用域
/*
scope 1 {
 debug v => _1; // in scope 1 at src/main.rs:3:5: 3:10
}
*/
// 2 非词法作用域 NLL: 案例 1

// Rust语言编译过程
// text -> tokens -> ast -> hir -> mir -> llvm ir -> llvm
// 在可变借用的作用域内不允许在开辟的子生命周期中执行可变借用
let mut v: Vec<i32> = vec![];
let vv = &v;

{
 // println!("{:?}", v[0]);
 // v.push(2); // 不允许可变借用
}

vv;
// 2 非词法作用域 NLL: 案例 2
// 替换问好
```

```

let s = "ab?c?d";

// 把字符串转成字符切片
let mut chars = s.chars().collect::<Vec<char>>();

println!("{:?}", chars);

for i in 0..s.len() {
 // 这里不可以用可变借用
 let mut words = ('a'..'z').into_iter();
 println!("{:?}", words);

 if chars[i] == '?' {
 // 获取左边和右边的字符
 let left = if i == 0 { None } else { Some(chars[i - 1]) };
 let right = if i == s.len() - 1 {
 None
 } else {
 Some(chars[i + 1])
 };

 // 在26个字母中寻找不等于左边也不等于右边的字母进行替换
 chars[i] = words
 .find(|&w| Some(w) != left && Some(w) != right)
 .unwrap();
 }
}

// 将字符收集转换为字符串
let s = chars.into_iter().collect::<String>();
println!("{:?}", s)

```

### 3.3.1.2 生命周期参数

1. 目的：为了避免出现悬垂指针
2. 晚限定与早限定

生命周期参数一般出现在函数参数的传递过程中以及自定义类型声明时

有两种方式：晚限定和早限定，早限定是一种更普遍的用法，尤其是实现trait或者关联函数时，不用在每个函数签名处声明生命周期参数

总结：

late bound：在具体调用时才自动生成具体的生命周期参数实例，不可以手动指定，编译器会检查本地变量

early bound：可以指定生命周期参数，会让编译器只检查参数类型以及生命周期参数，不检查本地变量

3. trait 对象中的生命周期参数

```
trait Foo<'a> {}
```

```

struct FooImpl<'a> {
 s: &'a [u32],
}
impl<'a> Foo<'a> for FooImpl<'a> {}

// trait 对象必须使用 Box 包裹
// 任何实现了某个 trait 的类型，它的实例都是 trait 对象
// trait 对象默认为静态生命周期，当作为返回值时，需要手动“缩短”（指定生命周期参数，如 'a)

// fn foo<'a, 'b: 'a>(s: &'a [u32]) -> Box<dyn Foo<'a> + 'a> { // 第一种写法
fn foo<'a>(s: &'a [u32]) -> Box<dyn Foo<'a> + 'a> {
 // 第二种写法
 Box::new(FooImpl { s: s })
}

```

#### 4. 高阶生命周期参数

```

use std::fmt::Debug;
trait DosSomething<T> {
 fn do_something(&self, value: T);
}

impl<'a, T: Debug> DosSomething<T> for &'a usize {
 fn do_something(&self, value: T) {
 println!("{}:?", value);
 }
}

// 高阶生命周期，高阶限定，for 语法，是一种 late bound
// fn foo<'a>(b: Box<dyn DosSomething<&'a usize>>) { 改动前
fn foo<'a>(b: Box<dyn for<'f> DosSomething<&'f usize>>) {
 // 不在当前作用域判断
 let s: usize = 10;
 b.do_something(&s) // 在 do something 函数作用域判断
}

let x = Box::new(&2usize);
foo(x)

```

注意：并不是所有的生命周期都是在当前作用域判断的

5. 闭包生命周期参数
6. trait 对象中的生命周期参数

### 3.3.2 类型系统

Rust编译器遵循类型理论：仿射类型：它是一种子结构类型系统。意义：资源最多只能被使用一次。

Rust类型系统有两种语义：移动语义（默认） 复制语义（该类型必须实现Copy trait：数据能够被安全的复制）

为什么实现Copy必须先实现Clone，它是编译器的行为，开发者再实现无用

```
pub trait Copy: Clone { }
```

哪些是移动语义？在运行时动态增长的类型，也就是说需要动态分配内存

Copy本质上是按位复制，并且不可以被重载，clone隐式调用，可以显式实现和调用

### 3.3.3 内存管理

1. 数据默认存储到栈上
2. 利用栈来自动管理堆内存（结合函数调用栈来理解，当栈帧被清除时，自动调用析构函数Drop，堆上的数据也被清空）

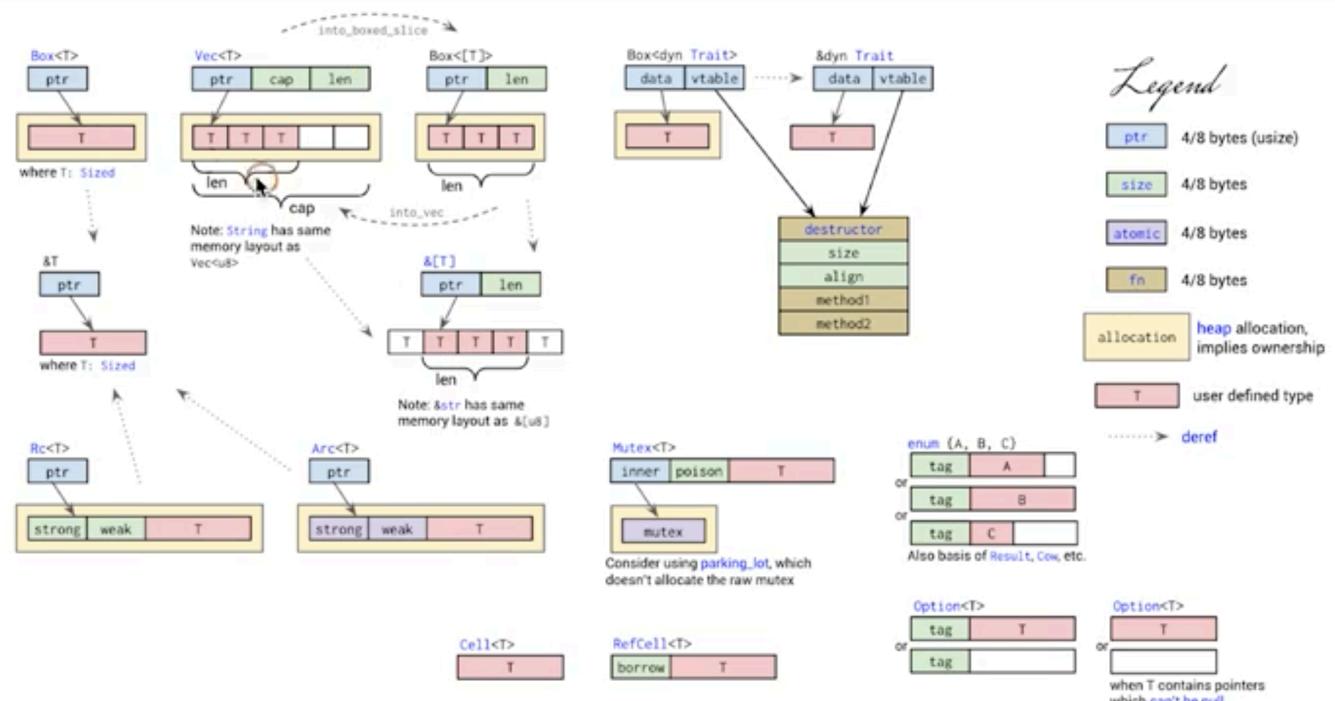
Box也叫做装箱类型，栈上会保留指针

Vec:确定性析构

Box: trait 对象在栈上，保留了数据指针和虚表指针

Rc和Arc引用计数的容器：可以共享所有权，强指针有所有权。锁和容器也类似（不是说强弱指针）

枚举：相当于每个枚举值前面都有tag



### 3.3.4 借用

借用本质上指的是所有权的借用。可以把它看作是一个指针（被借用者可以看作是内存位置），但是它是安全的，经过Rust编译器安全检查的。安全检查包括一些行为，比如可变与不可借用/使用等。Safe Rust中，引用永远是指向有效的数据

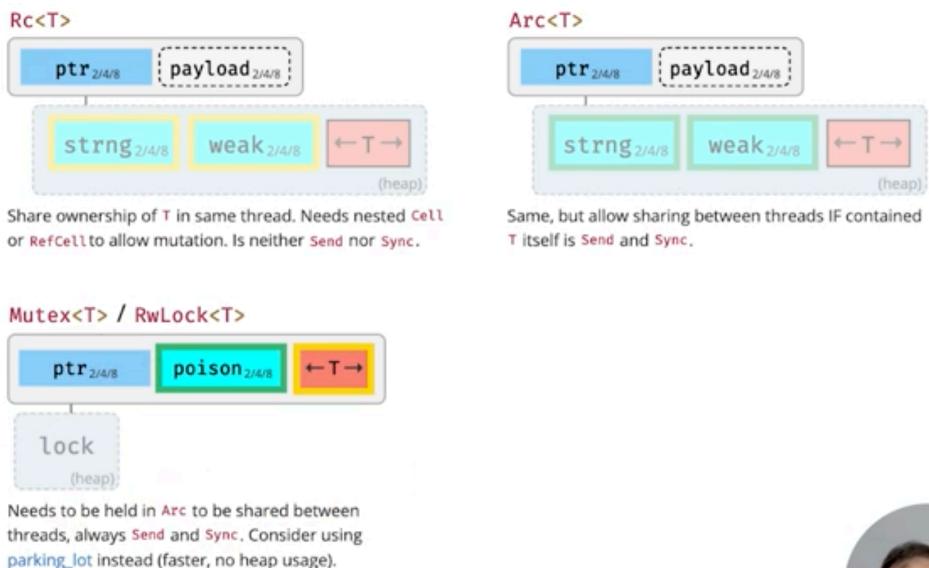
关于裸指针（没有安全的外衣）

### 3.3.5 共享

1. Rust中的Clone trait在语义上表示：所有权共享
2. 包含两种：一种是深拷贝，另一种是引用计数。但是二者共用一个clone trait

引用计数容器Rc和Arc以及同步所和互斥锁（Mutex和RwLock）

## 引用计数容器



## 多线程需要加锁

## 3.4 线程安全：线程和并发

### 3.4.1 本地线程

本地线程也叫内核线程，由操作系统来调度。

并发：同时应对很多事情的能力

并行：同时执行很多事情的能力

Rust使用了强大的类型系统以及两个专用的trait来在编译期时就发现并发安全问题

```
// Rust 中的线程

// 时间间隔
let duration = std::time::Duration::from_millis(30000);

println!("main thread ");
```

```

use std::thread;

// 使用 thread
let handle = thread::spawn(move || {
 println!("sub thread 1");

 let handle2 = thread::spawn(move || {
 println!("sub thread 2");
 thread::sleep(duration)
 });

 handle2.join().unwrap();
 thread::sleep(duration)
});

handle.join().unwrap();
thread::sleep(duration)

// rust 并不保证线程之间的引用之间的生命周期关系
// rust线程由操作系统调度

```

### 3.4.2 线程间共享数据

1. 手动实现必要的trait: 共享借用和所有权类型的数据
2. 使用第三方库 crossbeam: 共享借用 / 可变借用
3. 使用Arc和Mutex

```

// 在线程间共享数据

// 案例 1 通过借用检查, 消除数据竞争
use std::thread;
let mut v = vec![1, 2, 3, 4];
// thread::spawn(move || v.push(5)); // v 只能使用1次, 无法使用for 循环迭代加入元素

// 借用规则要求可变借用只能有一次, 避免了数据竞争 (多个线程同时使用 v)
// for i in 0..10 {
// thread::spawn(move || v.push(i));
// }

// 案例 2 通过函数来传递数据, 也不被允许
// 线程中没法传递引用, 因为不知道线程执行顺序
// 如果线程封装在函数中, 不知道函数会被在哪里调用以及调用多少次

// fn inner_func(vref: &mut Vec<u32>) {
// std::thread::spawn(move || vref.push(3));
// }

// 案例 4 只读也不能通过函数传递吗? 不能, 可能存在悬垂指针

```

```

// fn inner_func(vref: &Vec<u32>) {
// std::thread::spawn(move || println!("{}: {:?}", vref));
// }

// 案例 5 如何在线程间传递引用

// 5.1 不使用第三方库的实现 (加'static'类型)
use std::fmt;
struct Foo {
 string: String,
 v: Vec<f64>,
}

impl fmt::Display for Foo {
 fn fmt(&self, f: &mut fmt::Formatter<'_>) -> fmt::Result {
 write!(f, "{}: {:?}", self.string, self.v)
 }
}

// 封装到函数中
// 但是存在约束条件
fn test<T: Send + Sync + fmt::Display + 'static>(val: T) {
 thread::spawn(move || println!("{}: {:?}", val));
}

test("hello"); //&'static str
test(String::from("hello")); // String, 因为它是所有权的数据, 与程序同生同灭
test(5); // i32

// 内部的数据是由所有权的, 所以也可以作为参数传递
let foo = Foo {
 string: String::from("hello"),
 v: vec![1.2, 2.2, 3.2, 42.2],
};
test(foo);
// test(foo); 不能使用第二次

use std::time::Duration;
thread::sleep(Duration::new(1, 0));

// 5.2 使用第三方库crossbeam 的实现
// crossbeam::scope 共享数据
use crossbeam;
let mut vec = vec![1, 2, 3, 4, 5];

crossbeam::scope(|scope| {
 // scope 出来的子线程会在主线程关闭之前回收
 // 保证不会出现悬垂指针
})

```

```

scope.spawn(move |_|
 for e in &vec {
 println!("{:?}", e);
 }
);
}

.expect("a child thread panicked");

let mut v = vec![1, 2, 3, 4, 5];

crossbeam::scope(|scope| {
 // scope 出来的子线程会在主线程关闭之前回收
 // 不出现数据竞争
 for e in &mut v {
 scope.spawn(move |_| thread::sleep(Duration::from_secs(1)));
 }
})
.expect("a child thread panicked");

use std::sync::{Arc, Mutex};
// 5.3 也可以使用Arc和Mutex实现共享所有权
let v = Arc::new(Mutex::new(vec![1, 2, 3]));

// 每次都克隆一个
for i in 0..3 {
 let cloned_v = v.clone();
 thread::spawn(move || {
 cloned_v.lock().unwrap().push(i);
 });
}
}

```

### 3.4.3 构建无悔的并发系统

并发编程需要注意的三点：

原子性：保证操作是原子的

可见性：保证数据是同步的

顺序性：保证操作的顺序是正确的

并发编程的方式：

同步锁和无锁编程

锁带来的问题

性能：无锁编程可以最大化减少线程上下文切换、线程等待

死锁：引入无锁编程就不会产生死锁

无锁编程主要依靠原子类型，性能上并不总是优于锁编程

无锁编程和计算机组成密切相关：现代计算机一般都是多核三级缓存，带来缓存一致性问题；CPU指令重排；编译器指令重排。用内存屏障解决问题

内存屏障允许开发者在编写代码时在需要的地方加入它：内存屏障是指一种操作，它确保在该操作之前的内存访问完成，并且在该操作之后的内存访问不会在该操作之前执行。这有助于在多线程环境中维护内存的一致性和避免数据竞争。

CPU有四种屏障

内存模型：获取语义和释放语义

#### 1. 多线程并发

使用channel 和 condvar 模拟并行组件。Rust 只保证语言层面的安全，逻辑层面的安全并不保证

并发模型的最佳默认模式：事件循环 (event-loop)

### 3.4.4 无锁并发

#### 2. 无锁并发

原子类型：原子布尔值和数字，都提供了Ordering内存顺序：5种顺序，和LLVM以及C++20一致

原子类型还分硬件架构，ARM上的Linux没有原子类型

Rust提供了条件编译

内存顺序

```
pub enum Ordering {

 Relaxed, 原子类型只保证原子操作，不指定内存顺序（不指定内存屏障）
 Release, 当前线程内的所有写操作，对于其他对这个原子变量进行acquire得线程可见
 Acquire, 可以保证读到所有在Release之前发生的写入
 AcqRel, 对读取和写入施加acquire-release 语义，无法被重排
 SeqCst,
}
```

原子类型提供的方法：使用支持硬件的指令和方法

ABA问题

可以关注的库

### 3.5 trait 和泛型

让一个类型拥有方法有两种方式：自定义关联函数（使用其他类型作为泛型约束），为其实现trait

### 3.5.1 trait

#### trait四种作用

接口

类型标记

泛型限定

抽象类型

接口也是一种多态

作为泛型的限定

```
// trait 作为泛型限定
use std::string::ToString;

fn print<T: ToString>(v: T) {
 println!("{}", v.to_string());
}
```

抽象类型 (trait object) : 因为trait中包含了很多方法，在运行时都化作trait对象。用一个trait 对象可以表示同样实现了 trait的多种类型

trait 有两种分发类型：静态分发（单态化）：生成具体类型的函数

静态分发还有一种语法：impl trait

```
// 静态分发: impl trait

use std::fmt::Display;

// 返回一个实现了 Display trait 的类型
fn make_value<T: Display>(index: usize) -> impl Display {
 match index {
 0 => "Hello,World",
 1 => "Hello,world (1)",
 _ => panic!(),
 }
}

println!("{}", make_value::<&'static str>(0));
println!("{}", make_value::<&'static str>(1))
```

trait与生命周期

```

// trait 与生命周期
// fn make_debug<T>(_: T) -> impl std::fmt::Debug {
// 42u8
// }

// late bound
fn make_debug<'a, T: 'static>(_: &'a T) -> impl std::fmt::Debug {
 42u8
}

fn test() -> impl std::fmt::Debug {
 let value = "value".to_string();
 make_debug(&value)
}

```

### 3.5.2 trait 对象

是动态分发的一种

Any是Rust中仅有的一种自省机制，相当于反射机制。因为rust是编译型语言，所以作用有限，智能识别static（不能是引用类型），在运行时反射。

```

// 实现了Any trait 的类型到具体类型的转换
use std::fmt::Debug;

// 当函数参数是string时，可以转换为具体类型，否则什么都不干
fn log<T: Any + Debug>(value: &T) {
 let value_any = value as &dyn Any; // 先转为trait 对象
 match value_any.downcast_ref::<String>() {
 // 转为 String
 Some(as_string) => {
 println!("String ({}): {}", as_string.len(), as_string)
 }
 None => println!("{:?}", value),
 }
}

fn do_work<T: Any + Debug>(value: &T) {
 log(value)
}

let my_string = "hello world".to_string();
do_work(&my_string);
let my_i8 = 100;
do_work(&my_i8);

```

TypeId是全局唯一，当程序重新启动会发生变化

trait 对象：也是一组方法的集合

```
&dyn Trait or Box<dyn Trait>
```

```
use core::any::{Any, TypeId};
use std::sync::Arc;

// 模拟类
// 类的实例相当于trait 对象
struct Class {
 name: String,
 type_id: TypeId,
}

impl Class {
 fn new<T: 'static>() -> Self {
 Class {
 name: std::any::type_name::<T>().to_string(),
 type_id: TypeId::of::<T>(),
 }
 }
}

struct Instance {
 inner: Arc<dyn Any>, //相当于 Box<T>
}

impl Instance {
 fn new(obj: impl Any) -> Self {
 Self {
 inner: Arc::new(obj),
 }
 }

 fn instance_of(&self, class: &Class) -> bool {
 self.inner.as_ref().type_id() == class.type_id
 }
}

struct Foo {};
struct Bar {};

let foo_class = Class::new::<Foo>();
let bar_class = Class::new::<Bar>();

let foo_instance = Instance::new(Foo {});

assert!(foo_instance.instance_of(&foo_class));
assert!(!foo_instance.instance_of(&bar_class));
```

### 3.5.2.1 泛型和trait 对象实现模版方法

多个类型实现同一个trait

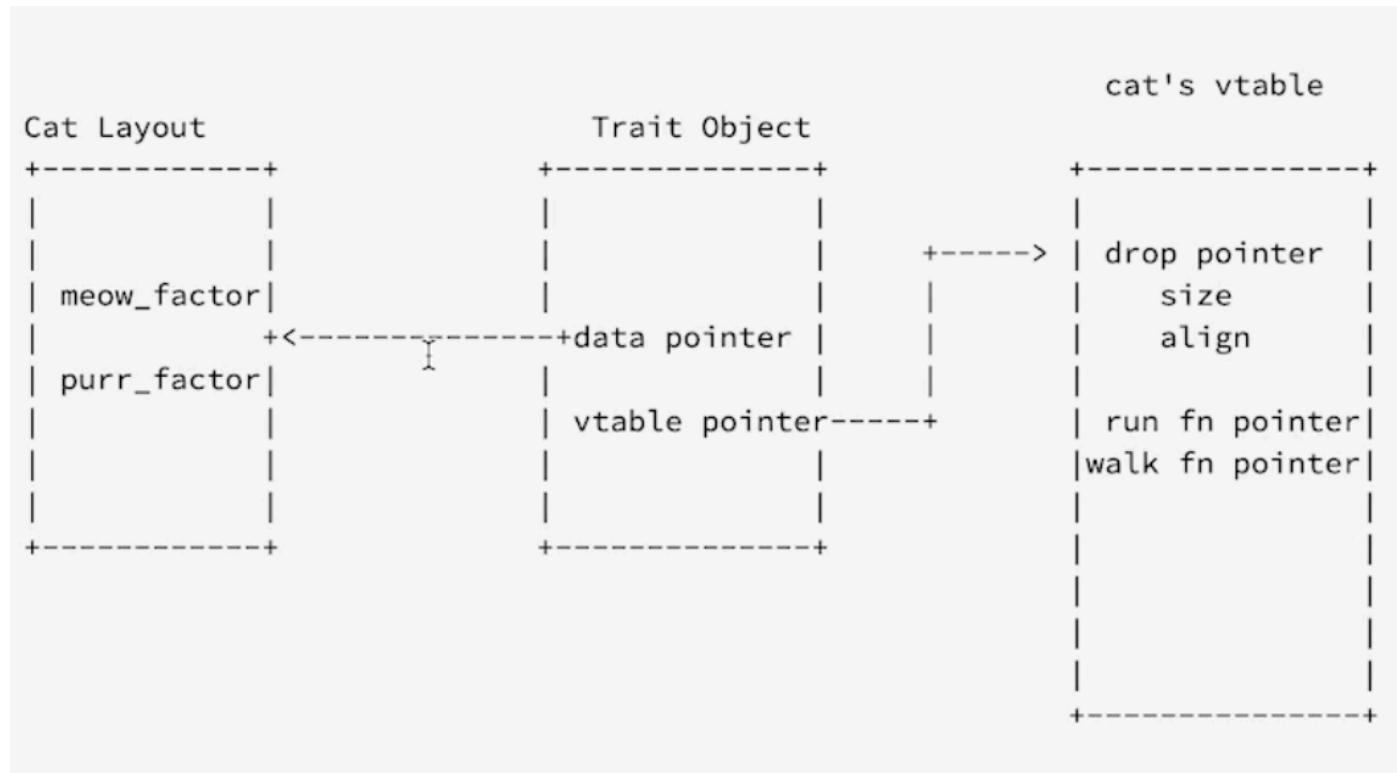
代表项目：actix-extras

### 3.5.2.2 trait对象的本质

trait定义了共同的行为

vtable存的是函数指针集

trait 对象本质上是一个虚表



### 3.5.2.3 trait 对象安全的本质

当多个类型实现了trait中的方法时，返回类型实例无法确定。对象安全就要确保方法能被安全的调用

编译器如何确保对象安全？如果trait能实现自己就是对象安全的

```
trait StarkFamily {
 fn last_name(&self) -> &'static str;
 fn totem(&self) -> &'static str;
}

trait TullyFamily {
 fn territory(&self) -> &'static str;
}

trait Children {
 fn new(first_name: &'static str) -> Self
}
```

```

where
 Self: Sized;

fn first_name(&self) -> &'static str;
}

impl StarkFamily for dyn Children {
 fn last_name(&self) -> &'static str {
 "Stark"
 }

 fn totem(&self) -> &'static str {
 "Wolf"
 }
}

impl TullyFamily for dyn Children {
 fn territory(&self) -> &'static str {
 "Riverrun City"
 }
}

struct People {
 first_name: &'static str,
}

impl Children for People {
 fn new(first_name: &'static str) -> Self
 where
 Self: Sized,
 {
 println!("hello,{:?} Stark", first_name);
 People {
 first_name: first_name,
 }
 }
 fn first_name(&self) -> &'static str {
 self.first_name
 }
}

fn fully_name(person: Box<dyn Children>) {
 println!(
 "--- Winter is coming, the lone {:?} dies, the packs lives ---",
 person.totem()
);
}

let full = format!("{} {}", person.first_name(), person.last_name());
println!("I'm {:?}", full);

```

```

 println!("My mother come from {:?}", person.territory());
}

let sansa = People::new("Sansa");
let aray = People::new("Aray");

let starks = Box::new(sansa);
fully_name(starks);

let starks = Box::new(aray);
fully_name(starks)

```

维护了两个虚表，safe\_table,nosafe\_vtable,where Self:sized,nosafe\_vtable

### 3.5.2.4 使用Enum 代替trait

当trait对象无法保证安全时的替代方案

trait 对象性能比较差，因为它在运行时，想要提高性能可以转为enum

```

// 类型不同，行为相同，通过trait实现
trait KnobControl {
 fn set_position(&mut self, value: f64);
 fn get_value(&self) -> f64;
}

struct LinearKnob {
 position: f64,
}

struct LogarithmicKnob {
 position: f64,
}

impl KnobControl for LinearKnob {
 fn set_position(&mut self, value: f64) {
 self.position = value
 }
 fn get_value(&self) -> f64 {
 self.position
 }
}

impl KnobControl for LogarithmicKnob {
 fn set_position(&mut self, value: f64) {
 self.position = value
 }

 fn get_value(&self) -> f64 {

```

```

 (self.position + 1.).log2()
 }
}

// 通过enum实现
// 将类型抽象到枚举体中

enum Knob {
 Linear(LinearKnob),
 Logarithmic(LogarithmicKnob),
}

impl KnobControl for Knob {
 fn set_position(&mut self, value: f64) {
 match self {
 Knob::Linear(inner_knob) => inner_knob.set_position(value),
 Knob::Logarithmic(inner_knob) => inner_knob.set_position(value),
 }
 }

 fn get_value(&self) -> f64 {
 match self {
 Knob::Linear(inner_knob) => inner_knob.get_value(),
 Knob::Logarithmic(inner_knob) => inner_knob.get_value(),
 }
 }
}

```

```

use core::ops::Add;
// 类型不同，行为相同，通过trait实现
trait KnobControl<T: Add + Add<Output = T> + Copy> {
 fn set_position(&mut self, value: T);
 fn get_value(&self, p: T) -> T;
}

struct LinearKnob<T: Add + Add<Output = T> + Copy> {
 position: T,
}

struct LogarithmicKnob<T: Add + Add<Output = T> + Copy> {
 position: T,
}

impl<T: Add + Add<Output = T> + Copy> KnobControl<T> for LinearKnob<T> {
 fn set_position(&mut self, value: T) {
 self.position = value
 }

 fn get_value(&self, p: T) -> T {

```

```

 self.position
 }
}

impl<T: Add + Add<Output = T> + Copy> KnobControl<T> for LogarithmicKnob<T> {
 fn set_position(&mut self, value: T) {
 self.position = value
 }

 fn get_value(&self, p: T) -> T {
 self.position + p
 }
}

// 通过enum实现
// 将类型抽象到枚举体中

enum Knob<T: Add + Add<Output = T> + Copy> {
 Linear(LinearKnob<T>),
 Logarithmic(LogarithmicKnob<T>),
}

impl<T: Add + Add<Output = T> + Copy> KnobControl<T> for Knob<T> {
 fn set_position(&mut self, value: T) {
 match self {
 Knob::Linear(inner_knob) => inner_knob.set_position(value),
 Knob::Logarithmic(inner_knob) => inner_knob.set_position(value),
 }
 }

 fn get_value(&self, value: T) -> T {
 match self {
 Knob::Linear(inner_knob) => inner_knob.get_value(value),
 Knob::Logarithmic(inner_knob) => inner_knob.get_value(value),
 }
 }
}

```

### 3.5.2.5 trait 覆盖实现

Rust trait中的方法不允许覆盖实现

但是可以使用trait 对象实现

### 3.5.2.6 trait 与 Self: Sized

什么时候需要用到。Rust中所有类型，默认都是T: Sized

```
// trait 中有默认实现时
// 并且默认实现的函数体中包含Self
```

```

trait WithConstructor {
 fn build(param: usize) -> Self
 where
 Self: Sized;
 fn new(param: usize) -> Self
 where
 Self: Sized,
 {
 Self::build(0)
 }

 fn t(&self);
}

struct A;

impl WithConstructor for A {
 fn t(&self) {
 println!("hello");
 }
 fn build(param: usize) -> Self
 where
 Self: Sized,
 {
 A
 }
}

let a = &A;
a.t()

```

### 3.5.2.7 trait 对象与Box

```

trait Test {
 fn foo(&self);

 fn works(self: Box<Self>) {
 println!("hello");
 }

 fn fails(self: Box<Self>)
 // where
 // Self: Sized, //限定了被调用,关闭; ? Sized 在类型声明时使用
 {
 self.foo();
 }
}

struct Concrete;

```

```

impl Concrete {
 fn hello(&self) {
 println!("hello");
 }
}

impl Test for Concrete {
 fn foo(&self) {
 ()
 }
 fn works(self: Box<Self>) {
 self.hello();
 }
 // 没有实现fails
}

let concrete: Box<dyn Test> = Box::new(Concrete);
// concretefails();
concrete.works();

```

## 3.6 Rust语言编程范式

Rust支持面向对象语言的一些特性，也支持函数式语言的特性。函数式style:

1. 默认不可变，但是rust可变
2. 支持递归，但rust不支持尾递归优化（推荐递归而不是优化）
3. 函数式一等公民，有限的高阶函数支持
4. 和类型/积类型

Rust语言式混合范式

### 3.6.1 面向编译器编程

洋葱模型：编译器->核->标准库->第三方库

## 3.7 Rust 错误处理

Rust是基于返回值的错误机制

Rust整体的错误机制

1. 类型系统保证函数契约
2. 断言用于防御
3. Option消除空指针失败
4. Result<T,E> 传播错误
5. Panic恐慌

## 3.7.1 消除失败

### 1. 类型系统保证函数契约

```
// 1 类型系统保证函数契约
fn sum(a: i32, b: i32) -> i32 {
 a + b
}

// sum(1u32, 2u32) 违反函数契约
```

### 2. 断言用于防御

```
// 2 断言用于防御

fn extend_vec(v: &mut Vec<i32>, i: i32) {
 assert!(v.len() == 5);
 v.push(i)
}

let mut vec = vec![1, 2, 3];
extend_vec(&mut vec, 4);
extend_vec(&mut vec, 5);
assert_eq!(5, vec[4]);
extend_vec(&mut vec, 6); // panic
```

## 3.7.2 错误处理：Option

分层错误处理：Option 有无，Result 对错

Option提供了一些方法可以方便操作，如map

```
// 3 Option
let maybe_some_string = Some(String::from("hello, world!"));
let maybe_some_len = maybe_some_string.map(|s| s.len());
assert_eq!(maybe_some_len, Some(13))
```

// 返回值类型都是Option可以使用链式调用，不需要一个个unwrap处理

```
fn double(val: f64) -> f64 {
 val * 2.
}

fn square(val: f64) -> f64 {
 val.powi(2 as i32)
}

fn inverse(val: f64) -> f64 {
```

```

 val * -1.
}

fn log(val: f64) -> Option<f64> {
 match val.log2() {
 x if x.is_normal() => Some(x),
 _ => None,
 }
}

fn sqrt(val: f64) -> Option<f64> {
 match val.sqrt() {
 x if x.is_normal() => Some(x),
 _ => None,
 }
}

let number = 20.;
let result = Option::from(number)
 .map(inverse)
 .map(double)
 .map(inverse)
 .and_then(log)
 .map(square)
 .and_then(sqrt);
match result {
 Some(x) => println!("x was {:?}", x),
 None => println!("this failed"),
}

```

map方法接受一个泛型参数，返回一个实现了FnOnce 闭包类型

### 3.7.3 Result

result的错误需要处理，当直接使用unwrap时，如果结果是Err，会发生Panic

什么样的Error才算，实现了Error trait，自定义必须实现该trait

引用第三方库

使用？

### 3.7.4 Panic

Panic的两种类型：unwinding（栈展开） aborting（中止）无法恢复

资源超过分配直接aborting

## 3.8 元编程

### 3.8.1 反射

Any: Rust中唯一反射，运行时反射

因为Rust是编译型语言，没有在运行时提供很多的反射功能。并且只有`'static`的类型才能支持动态运行时反射。

```
// case 1 反射
fn log<T: Any + Debug>(value: &T) {
 // 将具体类型转换为 trait 对象
 let value_any = value as &dyn Any;

 // 反射，判断类型，也叫自省
 match value_any.downcast_ref::<String>() {
 Some(as_string) => println!("string ({}): {}", as_string.len(), as_string),
 None => {
 println!("{:?}", value)
 }
 }
}

fn do_work<T: Any + Debug>(value: &T) {
 log(value)
}

let my_string = "hello world".to_string();
do_work(&my_string);
let my_i8: i8 = 100;
do_work(&my_i8);

// 反射如何实现
// pub trait Any: 'static {
// pub fn type_id(&self) -> TypeId;
// }

// 为 dyn Any 实现了 fn is<T: Any>(&self) -> bool; 方法
// 他也是线程安全的

// case 2

use std::any::Any;

trait Foo: Any {
 fn as_any(&self) -> &dyn Any;
}

impl<T: Any> Foo for T {
 fn as_any(&self) -> &dyn Any {
 self
```

```
}

#[derive(Debug)]
struct Bar {}

#[derive(Debug)]
struct Baz {}

impl PartialEq for dyn Foo {
 fn eq(&self, other: &dyn Foo) -> bool {
 let me = self.as_any();
 let you = other.as_any();

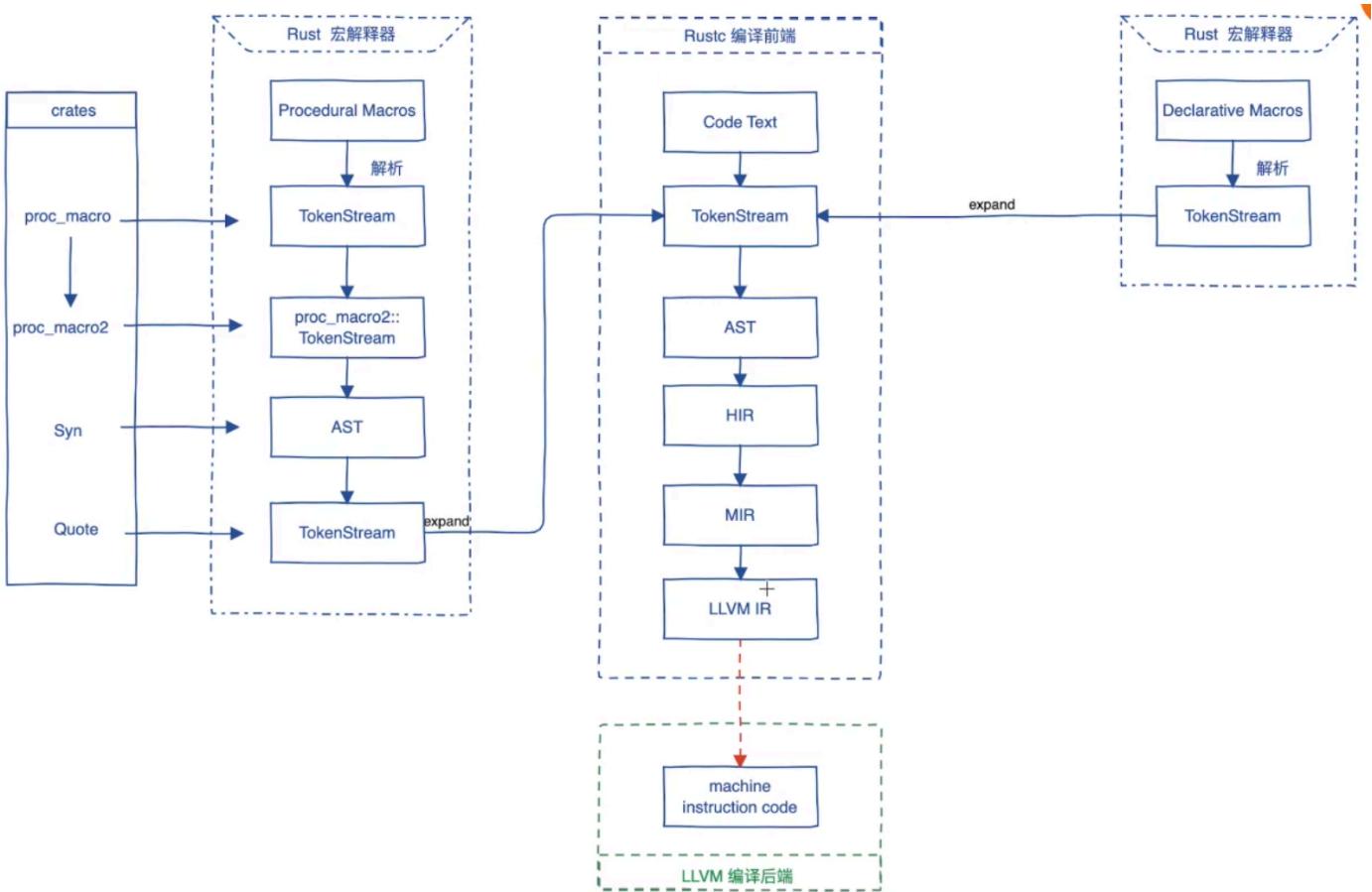
 if me.is::<Baz>() && you.is::<Baz>() {
 true
 } else {
 false
 }
 }
}

let bar = Bar {};
let baz = Baz {};
let foo1 = &bar;
let foo2 = &baz;

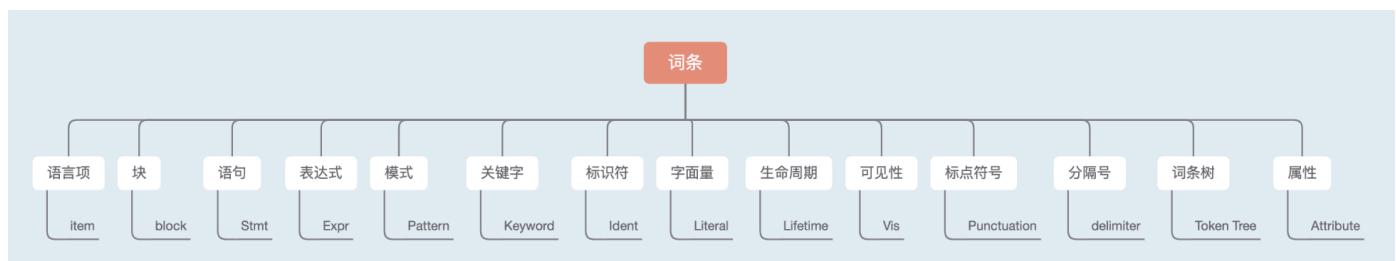
println!("{:?}", foo1);
println!("{:?}", foo2);
```

## 2.8.2 宏

宏是代码生成的一种技术，在此之前需要先理解rust编译过程



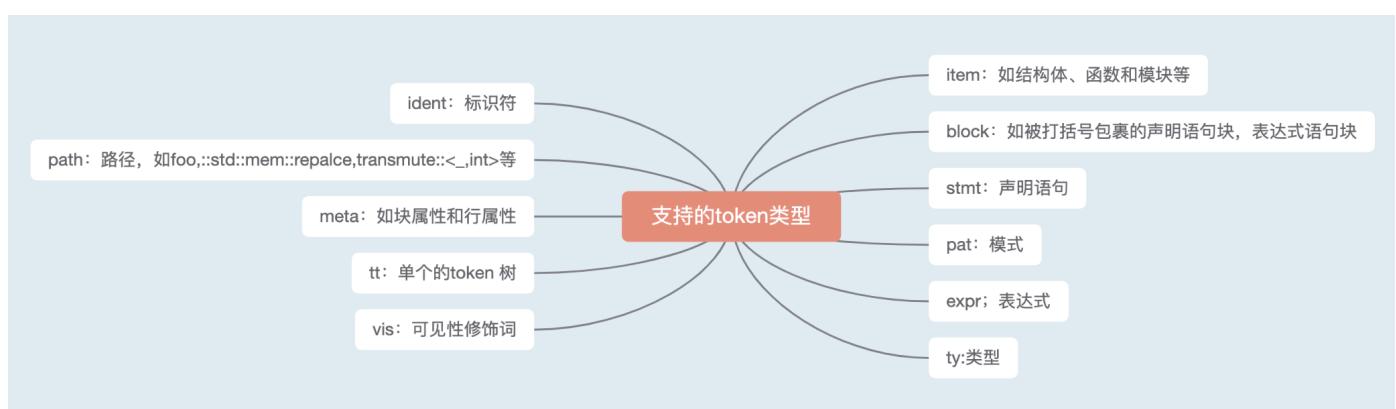
元编程也叫DSL，Domain Special Language



### 2.8.2.1 声明宏

把宏展开为TokenStream。只做替换几乎不做计算。如果是Token匹配，就是声明宏。

声明宏也是在分词阶段进行正则表达式的一种匹配



```
// 声明宏: 传入两个表达式, 正则匹配表达式
```

```

// 入参
// 替换
macro_rules! unless {
 ($arg:expr => $branch:expr) => {
 // 自定义自己的语言
 if !$arg {
 $branch
 }
 };
}

fn cmp(a: i32, b: i32) {
 unless!(a > b => println!("{} < {}", a, b))
}

let (a, b) = (1, 3);
cmp(a, b)

```

## 2.8.2.2 过程宏

### 1. derive 宏

更加复杂，在Tokenstream上又构建了自己的AST，为了更强大的计算

以serde库为例

过程宏有三种：一种是类似于声明宏那样的函数调用，第二种：派生宏，第三种：属性宏

派生宏原理：把结构体解析为词条流，使用宏派生宏陪里面专门定义的词条处理的方法，然后结合自定义的AST来处理

```

#[derive(Serialize, Deserialize)]
#[serde(deny_unknown_fields)]
struct S {
 #[serde(default)]
 f: i32,
}

```

如何实现过程宏？离不开过程宏三件套：syn(ast) quote(ast转为词条流) proc-macro2

proc-macro2 库：使用仅限于过程宏

syn（依赖于proc-macro2）和quote：配合使用，syn是把proc-macro2的TokenStream 转为AST，quote是再转回去。相当于这两个库配合又再加了一层

syn提供了一些数据结构：其实是语法树

过程宏的实例：Bang宏实现原理

宏一般独立一个crate，几乎可以做任何事情》宏代码调试工具：darling，可以在宏代码打log，cargo expand展开查看错误

第三方有哪些好用的宏代码

Derive-new 和 derive-more

过程宏的逻辑：解析->匹配模版->组装模版->输出为TokenStream

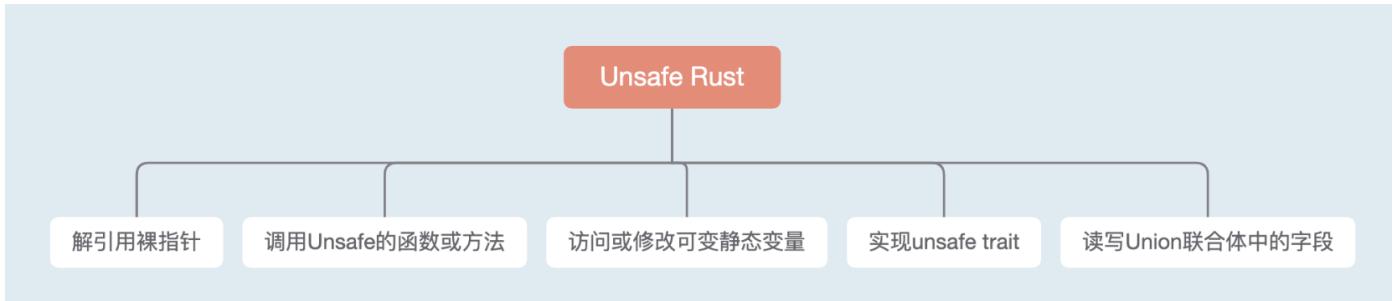
## 2. 属性宏

语法相对来说更加自由：案例：log-derive;rocket

## 3.9 Unsafe Rust

是Rust的超集，Unsafe rust也是有安全检查的

以下几种情况Rust不会提供任何安全检查



解引用裸指针：`*const T` 和 `*mut T` 两种指针类型，因为其和C语言中的指针十分相近，所以叫原生指针

原生指针的特点：

1. 不保证指向合法内存，如空指针
2. 不能像智能指针那样，自动清理内存
3. 没有生命周期的概念，编译器不会对其进行借用检查
4. 不能保证线程安全

```
// 解引用静态变量
static mut COUNTER: u32 = 0;
let inc = 3;

unsafe {
 COUNTER += inc;
 println!("Counter: {}", COUNTER);
}
```

safe rust构建于unsafe rust之上，凭什么safe？

官方保证：1. unsafe 在调用时注明安全边界；2. 实现了形式化验证；3. 安全数据库

## 3.9.1 安全抽象

从指针到引用。从不安全抽象为安全

### **3.9.2 Drop检查**

### **3.9.3 型变**

协变和逆变

### **3.9.4 NonNull和**

NonNull 协变

## **4 异步编程**

---

### **4.1 异步I/O模型**

---

进行等待返回，同步，否则异步

#### **4.1.1 同步和异步**

等待数据准备好这段时间

调用立刻返回数据/错误，然后轮询，非阻塞

现在常用的是IO多路复用：同步：数据等待和数据拷贝，第二阶段永远阻塞。多路复用也是一种同步IO模型。实现一个线程监视多个文件句柄

### **4.2 epoll 和 io\_uring**

---

epoll: 是一个同步的多路复用，实际上是一种事件通知机制：三个函数。两种触发机制。惊群问题已经解决

io\_uring: 用户态和核心共享两个环形内存

### **4.3 事件驱动编程模型**

---

处理IO复用的编程模型相当复杂，为了简化编程，提出了反应器模式和主动器模式

### **4.4 epoll代码实践**

---

使用三个系统调用函数。安卓，Linux都用

### **4.5 Reactor 代码实践**

---

事件驱动编程模型。读写，注册事件

### **4.6 MiniMio代码实践**

---

跨平台抽象，mio库

Linux和win有不同的系统抽象。抽象一个selector去选择不同的平台

### **4.7 Mio代码实践**

---

## 4.7.1 epoll接口

它是一个生产环境下的库

tcp

udp: poll::new 系统调用；轮询；建立UDP链接；处理等

## 4.7.2 其他代码

Waker;唤醒

io\_source 实现了Source trait

对不同的平台底层进行抽象

## 4.8 异步编程模型

---

与其他语言相比的特点：

1. Rust只提供零成本的异步编程抽象而不内置运行时，运行时可以替换如tokio
2. 基于Generator实现的future，在future的基础上提供 async/await语法糖，本质是一个状态机
3. Node.js依赖于V8，Go内置了运行时

为什么需要异步？

1. 对极致性能的追求
2. 对编程体验的追求

异步编程模型的发展阶段

1. callback (回掉地狱)
2. Promise/Future (会产生很多内嵌Future)
3. async/await：拥有了和同步代码的一致体验

异步任务可以看作是一种绿色线程

Future代表异步计算

## 4.8.1 Future

实现了该trait就可以异步计算：第三方库：future-rs

## 4.8.2 编写异步echo服务

1. 建立tcp链接
2. 处理tcp 流：read /write
3. poll/select epoll

## 4.8.3 异步Task模型

调度线程中的协程就是运行时。

## 4.8.4 Waker实现

一个task可以看作是一个线程中的微线程

## 4.9 异步库源码导读

异步运行时的实现机制：Future channel 是task之间通讯

Pin异步运行时中相当于一个模版

Future 的流相当于异步迭代器

Future task

## 4.10 async-await 语法

async的两种用法：`async fn` 和 `async {}`

Await 将暂停函数执行。如果用锁的话尽量使用Future提供的锁

```
use std::future::Future;
// async 真正会返回 Future<Output = u8>, 而不是看上去的u8
async fn foo() -> u8 {
 5 // 去糖后是Future
}

// async 块用法, 返回 "impl Future<Output = u8>""
fn bar() -> impl Future<Output = u8> {
 async { // 块返回值
 let x = foo().await;
 x + 5
 }
}
```

### 4.10.1 生成器

async / await 对应底层生成器为 resume/yield。yield是暂停点。和闭包的区别在于能暂停，底层实际上是一个状态机。和闭包底层也非常相似。

4.10.2 Rust

Rust解决自引用

为什么要用Pin?

## 4.10.2 Pin与Unpin

是一种使用类型系统的解决方案。Pin防止得到可变借用乱用

## 4.11 no\_std异步生态

---

核心库一般是使用在wasm和嵌入式，这些场景一般没有堆分配。所以关于堆分配的一些集合找不到运行时

1. async-std 异步的，专门处理异步io
2. tokio 最成熟的，生产级应用比较多
3. smol-rs 轻量的运行时 封装了很多底层的库
4. gloomio
5. bastion 目标是高可用的

## 4.12 实现异步缓存

---

Rust中的B树命中率更高。异步过程中构建组件等也是异步的。

多线程或者异步使用B树需要加锁,smol 实现了一些锁和屏障.B树和HashMap使用一样的。

区分同步和异步代码

如何清理过期缓存? redis: 按照频率, 定期删除策略

# 5 Rust异步Web框架

---

## 5.1 Rocket

---

充分的利用了

# 6 知名Rust项目

---

代码组织方式: Rust推荐整个项目使用多个crate构建

## 6.1 Rust

---

主要是编译器的实现

## 6.2 Wasmtime

---

字节码联盟维护的一个JIT的WebAssembly运行时，使用的编译器是Cranelift

## 6.3 Futures-rs

---

官方提供的一个运行时实现

## 6.4 async-std

---

标准库对async的实现

## 6.5 Tokio

---

比较成熟的异步运行时

## 6.6 Rocket

---

Web框架

## 6.7 Actix-web

---

Web框架

## 6.8 TiKV

---

数据库