## **Stacked Borrows:** An Aliasing Model for Rust (Work in Progress)

**Ralf Jung**, Hai Dang, and Derek Dreyer MPI-SWS, Germany

WG2.8, Bordeaux, May 2019



### A safe & modern systems PL



### A safe & modern systems PL

- First-class functions
- Polymorphism/generics
- Algebraic datatypes
- Traits  $\approx$  Type classes



## A safe & modern systems PL

- First-class functions
- Polymorphism/generics
- Algebraic datatypes
- Traits  $\approx$  Type classes



• Control over resource management (e.g., memory allocation and data layout)

## A safe & modern systems PL

- First-class functions
- Polymorphism/generics
- Algebraic datatypes
- Traits  $\approx$  Type classes



- Control over resource management (e.g., memory allocation and data layout)
- Strong type system guarantees:
  - Type & memory safety; data-race freedom





# Rust type system

# Rust type system



- 1. Full ownership: T
  - + mutation, deallocation
  - aliasing



- 2. Mutable reference: &mut T (borrowed)
  - + mutation
  - aliasing, deallocation
- 3. Shared reference: &T (borrowed) + aliasing
  - mutation, deallocation

Rust enforces this via ownership & borrowing: 1. Full ownership: T + mutation, deallocation

Rust's reference types provide strong aliasing information.

The optimizer should exploit that!

3. Shared reference: &T
 (borrowed)
 + aliasing
 - mutation, deallocation

```
fn test_unique(x: &mut i32) -> i32 {
 *x = 42;
 // unknown_function_1 cannot have an alias to x
 unknown_function_1();
 return *x; // must return 42
}
```

fn test\_noalias(x: &mut i32, y: &mut i32) -> i32 {
 // x, y cannot alias: they are unique pointers
 \*x = 42;
 \*y = 37;
 return \*x; // must return 42
}

```
fn test_shared(x: &i32) -> bool {
   let val = *x;
   // unknown_function_2 cannot mutate x
   unknown_function_2(x);
   return *x == val; // must return true
}
```

Eull ownorchine T

Rust's reference types provide strong aliasing information.

The optimizer should exploit that.

+ aliasing - mutation, deallocation

Full ownorship. T

Rust's reference types provide strong aliasing information.

The optimizer should exploit that. But there is a problem:

> + aliasing - mutation, deallocation

Eull ownorchin. T

Rust's reference types provide strong aliasing information.

The optimizer should exploit that.

But there is a problem:

### **UNSAFE CODE!**

+ aliasing - mutation, deallocation

# Unsafe code can access hazardous operations that are banned in safe code:

unsafe fn hazardous(x: usize) -> i32 {
 // \*const T is the type of raw (unsafe) pointers
 let x\_ptr = x as \*const i32;
 return \*x\_ptr; // dereferencing an arbitrary integer
}

- Used for better performance, FFI, implementing many standard library types
- Generally encapsulated by safe APIs

- 11: fn test\_unique(x: &mut i32) -> i32 {
- 12: \*x = 42;
- 13: unknown\_function\_1();
- 14: return \*x;
- 15: }

1: static mut ALIAS: \*mut i32 = std::ptr::null\_mut();

```
2: fn main() {
```

```
3: let 1 = &mut 0;
```

```
4: unsafe { ALIAS = 1 as *mut i32; }
```

```
5: println!("The answer is {}", test_unique(&mut *1));
```

6: // prints: The answer is 7

```
7: }
```

```
8: fn unknown_function_1() {
```

```
9: unsafe { *ALIAS = 7; }
```

```
10: }
```

```
11: fn test_unique(x: &mut i32) -> i32 {
```

```
12: *x = 42;
```

```
13: unknown_function_1();
```

```
14: return *x;
```

```
15: }
```





Work-in-progress aliasing model defining which pointers may be used to access memory, ensuring:

- Uniqueness of mutable references
- Immutability of shared references

Implemented in Miri, an experimental interpreter for Rust's MIR

https://github.com/rust-lang/miri

Already uncovered 9 bugs in Rust's standard library!

Work-in-progress aliasing model defining which pointers may be used to access memory, ensuring:

- Uniqueness of mutable references
- Immutability of shared references

Work-in-progress aliasing model defining which pointers may be used to access memory, ensuring:

- Uniqueness of mutable references
- Immutability of shared references

- 1: let mut 1 = 0;
- 2: let a = &mut 1;

- 1: let mut l = 0;
- 2: let a = &mut 1;
- 3: let b = &mut \*a;

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &mut *a;
4: *b = 3;
```

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &mut *a;
4: *b = 3;
5: *a = 4;
```

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &mut *a;
4: *b = 3;
5: *a = 4;
6: *b = 4; // ERROR: lifetime of 'b' has ended
```

1: let mut l = 0; 2: let a = &mut l; 3: let b = &mut \*a; 4: \*b = 3; 5: \*a = 4; 6: \*b = 4; // ERROR: lifetime of 'b' has ended 1: let mut l = 0; 2: let a = &mut l; 3: let b = &mut \*a; 4: \*b = 3; 5: \*a = 4; 6: \*b = 4; // ERROR: lifetime of 'b' has ended

- Chain of borrows:
  - ${\tt l}$  borrowed to  ${\tt a}$  reborrowed to  ${\tt b}$
- Well-bracketed: no ABAB



- 1 borrowed to a reborrowed to b
- Well-bracketed: no ABAB

### **Stacked Borrows ingredients**



Pointer values carry a tag: (*PtrVal* :=  $Loc \times \mathbb{N}_{\perp}$ ) 0x40[1], 0x41[ $\perp$ ]

Every location in memory comes with an associated stack: (Mem := Loc  $\stackrel{\text{fin}}{\longrightarrow}$  Byte  $\times$  Stack)

0x40: 0xFE, [0: Unique, 1: Unique] 0x41: 0xFE, [0: Unique,  $\perp$ : SharedRW]

:
#### **Stacked Borrows ingredients**



#### **Stacked Borrows ingredients**



1: static mut ALIAS: \*mut i32 = std::ptr::null\_mut();

```
2: fn main() {
```

```
3: let 1 = &mut 0;
```

```
4: unsafe { ALIAS = 1 as *mut i32; }
```

```
5: println!("The answer is {}", test_unique(&mut *1));
```

6: // prints: The answer is 7

```
7: }
```

```
8: fn unknown_function_1() {
```

```
9: unsafe { *ALIAS = 7; }
```

```
10: }
```

```
11: fn test_unique(x: &mut i32) -> i32 {
```

```
12: *x = 42;
```

```
13: unknown_function_1();
```

```
14: return *x;
```

```
15: }
```

```
1: let 1 = &mut 0;
```

- 2: let ALIAS = 1 as \*mut i32;
- 3: let x = &mut \*1;

```
4: *_{X} = 42;
```

- 5: unsafe { \*ALIAS = 7; }
- 6: println!("The answer is {}", \*x);

1: let 1 = &mut 0; // Tag: 0

1: let 1 = &mut 0; // Tag: 0

Stack: [0: Unique] 1: let l = &mut 0; // Tag: 0 2: let ALIAS = l as \*mut i32; // Tag: ⊥

Stack:
[0: Unique, ⊥: SharedRW]

Find permission for old tag 0 on stack; add new permission  $\perp$ : SharedRW above it 1: let l = &mut 0; // Tag: 0
2: let ALIAS = l as \*mut i32; // Tag: ⊥
3: let x = &mut \*l; // Tag: 1

Stack: [0: Unique, 1: Unique]

Find permission for old tag 0 on stack; remove incompatible  $\perp$ : SharedRW above; push new permission 1: Unique 1: let l = &mut 0; // Tag: 0
2: let ALIAS = l as \*mut i32; // Tag: ⊥
3: let x = &mut \*l; // Tag: 1
4: \*x = 42:

## Stack: [0: Unique, 1: Unique]

Find permission for tag 1 on stack; remove incompatible items above (none)

```
1: let l = &mut 0; // Tag: 0
2: let ALIAS = l as *mut i32; // Tag: ⊥
3: let x = &mut *l; // Tag: 1
4: *x = 42;
5: unsafe { *ALIAS = 7; }
```

Stack:
[0: Unique, 1: Unique]

Find permission for tag  $\perp$  on stack – there is no such item!

```
1: let 1 = &mut 0; // Tag: 0
```

- 2: let ALIAS = 1 as \*mut i32; // Tag:  $\perp$
- 3: let x = &mut \*1; // Tag: 1

```
4: *x = 42;
```

5: unsafe { \*ALIAS = 7; }

# It is undefined behavior to use a pointer whose tag is not on the stack.

Find permission for tag  $\perp$  on stack – there is no such item!

#### Stacked Borrows rules (so far)

- Memory access: find permission for our tag, remove incompatible items above
- Assigning fresh tags:
  - Taking a reference (&mut term): fresh tag n. find old tag; remove incompatible above; push new
  - Reference to raw pointer (term as \*mut T): tag ⊥.
     find old tag; add new just above

```
fn test_unique(x: &mut i32) -> i32 {
 *x = 42;
 // unknown_function_1 cannot have an alias to x
 unknown_function_1();
 return *x; // must return 42
}
```







x's tag with Unique permission is at the top of the stack

fn test\_unique(x: &mut i32) -> i32 {

We assumed that unknown\_function\_1 cannot have a reference with x's tag!

UB unless X's permission is still in the stack If unknown\_function\_1 accesses this memory, it will pop x's permission off the stack

```
fn test_unique(x: &mut i32) -> i32 {
  retag(x); // think: x = &mut *x;
  *x = 42;
  // unknown_function_1 cannot have an alias to x
  unknown_function_1();
  return *x; // must return 42
}
```

x gets a fresh tag with Unique permission pushed to top of the stack

fn test\_unique(x: &mut i32) -> i32 {
 retag(x); // think: x = &mut \*x;
 \*x = 42;
 // unknown\_function\_1 cannot have an alias to x
 unknown\_function\_1();
 return \*x; // must return 42
}

}

x gets a fresh tag with Unique permission pushed to top of the stack

fn test\_unique(x: &mut i32) -> i32 {
 retag(x); // think: x = &mut \*x;
 \*x = 42;
 // unknown\_function\_1 cannot have an alias to x
 unknown\_function\_1();
 return \*x; // must return 42

unknown\_function\_1 cannot guess or forge our tag: if it accesses this memory, it will pop x's tag off the stack

x gets a fresh tag with Unique permission pushed to top of the stack

fn test\_unique(x: &mut i32) -> i32 {
 retag(x); // think: x = &mut \*x;
 \*x = 42;
 // unknown\_function\_1 cannot have an alias to x
 unknown\_function\_1();

return \*x; // must return 42

UB unless X's permission is still in the stack

}

unknown\_function\_1 cannot guess or forge our tag: if it accesses this memory, it will pop x's tag off the stack

#### Stacked Borrows rules (so far)

- Memory access: find permission for our tag, remove incompatible items above
- Assigning fresh tags:
  - Taking a reference (&mut term): fresh tag n. find old tag; remove incompatible above; push new
  - Reference to raw pointer (term as \*mut T): tag ⊥.
     find old tag; add new just above

#### Stacked Borrows rules (so far)

- Memory access: find permission for our tag, remove incompatible items above
- Assigning fresh tags:
  - Taking a reference (&mut term): fresh tag n. find old tag; remove incompatible above; push new
  - Reference to raw pointer (term as \*mut T): tag ⊥.
     find old tag; add new just above

• Retag when reference "enters" function (argument, load, call returns)

### **Stacked Borrows**

Work-in-progress aliasing model defining which pointers may be used to access memory, ensuring:

- Uniqueness of mutable references
- Immutability of shared references

```
10:
11: fn test_shared(x: &i32) -> bool {
12: let val = *x;
13: unknown_function_2(x);
14: return *x == val;
15: }
```

```
1: fn main() {
```

- 2: let 1 = &mut 0;
- 3: println!("Test result: {}", test\_shared(&\*1));

```
4: // prints: Test result: false
```

- 5: }
- 6: fn unknown\_function\_2(x: &i32) {

```
7: let ptr = x as *const i32 as *mut i32;
```

```
8: unsafe { *ptr = 7; }
```

```
9: }
```

```
10:
```

```
11: fn test_shared(x: &i32) -> bool {
```

```
12: let val = *x;
```

```
13: unknown_function_2(x);
```

```
14: return *x == val;
```

```
15: }
```

- 1: fn main() {
- 2: let 1 = &mut 0;
- 3: println!("Test result: {}", test\_shared(&\*l));
- 4: // prints: Test result: false
- 5: }
- 6: fn unknown\_function\_2(x: &i32) {
- 7: let ptr = x as \*const i32 as \*mut i32;
- 8: unsafe { \*ptr = 7; } 9: }

Writes into x after some pointer casts

```
11: fn test_shared(x: \&i32) -> bool {
```

```
12: let val = *x;
```

- 13: unknown\_function\_2(x);
- 14: return \*x == val;

15: }

10:

- 1: fn main() {
- 2: let 1 = &mut 0;
- 3: println!("Test result: {}", test\_shared(&\*l));
- 4: // prints: Test result: false
- 5: }
- 6: fn unknown\_function\_2(x: &i32) {
- 7: let ptr = x as \*const i32 as \*mut i32;
- 8: unsafe { \*ptr = 7; }

9: }

10:

Writes into x after some pointer casts

11: fn test\_shared(x: &i32) -> bool {

```
12: let val = *x;
```

13: unknown\_function\_2(x);

Overwrites **\***x with 7

14: return \*x == val;

15: }

- 1: fn main() {
- 2: let 1 = &mut 0;
- 3: println!("Test result: {}", test\_shared(&\*l));
- 4: // prints: Test result: false

fn Mutating \*x inside
 unknown\_function\_2 must be UB
 to justify the optimization.

11: **fn** test\_shared(x: **&i32**) -> bool {

```
12: let val = *x;
```

- 13: unknown\_function\_2(x);
- 14: return \*x == val;

15: }

- 1: let mut l = 0;
- 2: let a = &mut 1;

- 1: let mut l = 0; 2: let a = &mut l;
- 3: let b = &\*a;

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &*a;
4: let _val = *b;
```

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &*a;
4: let _val = *b;
5: let _val = *a;
6: let _val = *b;
```

```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &*a;
4: let _val = *b;
5: let _val = *a;
6: let _val = *b;
7: *a = 1;
8: let _val = *b; // ERROR: lifetime of 'b' has ended
```

1: let mut l = 0; 2: let a = &mut l; 3: let b = &\*a; 4: let \_val = \*b; 5: let \_val = \*a; 6: let \_val = \*b; 7: \*a = 1; 8: let \_val = \*b; // ERROR: lifetime of 'b' has ended 1: let mut l = 0; 2: let a = &mut l; 3: let b = &\*a; 4: let \_val = \*b; 5: let \_val = \*a; 6: let \_val = \*b; 7: \*a = 1; 8: let \_val = \*b; // ERROR: lifetime of 'b' has ended

- Reads allowed through a and b...
- ... until first write through a
- No mutation between creation and use of a shared reference
```
1: let mut l = 0;
2: let a = &mut l;
3: let b = &*a;
4: let _val = *b;
5: let _val = *a;
```

### Shared references allow reads but no writes with other pointers.

• Reads allowed through a and b...

- ... until first write through a
- No mutation between creation and use of a shared reference

ended

### **Stacked Borrows ingredients**



#### **Stacked Borrows ingredients**

Pointers carry a tag: (PtrVal :=  $Loc \times \mathbb{N}_{\perp}$ ) 0x40[1], 0x41[ $\perp$ ]

Every location in memory comes with an associated stack: (*Mem* :=  $Loc \stackrel{\text{fin}}{\longrightarrow} Byte \times Stack$ )

0x40: 0xFE, [0: Unique, 1: Unique] 0x41: 0xFE, [0: Unique,  $\perp$ : SharedRW] 0x42: 0x00, [0: Unique, 1: SharedRO]

:

:

23

#### **Stacked Borrows ingredients**

:

Pointers carry a tag: (PtrVal :=  $Loc \times \mathbb{N}_{\perp}$ ) 0x40[1], 0x41[ $\perp$ ]

Every location in memory comes with an associated stack: (*Mem* :=  $Loc \stackrel{\text{fin}}{\longrightarrow} Byte \times Stack$ )

0x40: 0xFE, [0: Unique, 1: Unique] 0x41: 0xFE, [0: Unique,  $\perp$ : SharedRW] 0x42: 0x00, [0: Unique, 1: SharedRO]

Can be read but not written by 1.

```
1: fn main() {
```

- 2: let 1 = &mut 0;
- 3: println!("Test result: {}", test\_shared(&\*1));

```
4: // prints: Test result: false
```

```
5: }
```

6: fn unknown\_function\_2(x: &i32) {

```
7: let ptr = x as *const i32 as *mut i32;
```

```
8: unsafe { *ptr = 7; }
```

```
9: }
```

```
10:
```

```
11: fn test_shared(x: &i32) -> bool {
```

```
12: let val = *x;
```

```
13: unknown_function_2(x);
```

```
14: return *x == val;
```

```
15: }
```

1: let l = &mut 0; 2: let x = &\*l; 3: let val = \*x; 4: let ptr = x as \*const i32 as \*mut i32; 5: unsafe { \*ptr = 7; } 6: let test = \*x;

7: println!("Test result: {}", test == val);

1: let 1 = &mut 0; // Tag: 0

Stack: [0: Unique] 1: let l = &mut 0; // Tag: 0 2: let x = &\*l; // Tag: 1

## Stack: [0: Unique; 1: SharedRO]

Find read permission for 0 on stack; remove read-incompatible items above (none); push new permission 1: SharedRO 1: let l = &mut 0; // Tag: 0
2: let x = &\*l; // Tag: 1
3: let val = \*x;

### Stack: [0: Unique; 1: SharedRO]

Find read permission for 1 on stack; remove read-incompatible items above (none)

1: let l = &mut 0; // Tag: 0
2: let x = &\*l; // Tag: 1
3: let val = \*x;
4: let ptr = x as \*const i32 as \*mut i32; // Tag: ⊥

## Stack: [0: Unique; 1: SharedRO; ⊥: SharedRO]

Find read permission for 1 on stack; remove read-incompatible items above (none); push new permission  $\perp$ : <u>SharedRO</u> 1: let l = &mut 0; // Tag: 0
2: let x = &\*l; // Tag: 1
3: let val = \*x;
4: let ptr = x as \*const i32 as \*mut i32; // Tag: ⊥
5: unsafe { \*ptr = 7; }

# Stack: [0: Unique; 1: SharedR0; L: SharedR0]

ptr cannot be used for writing: ⊥only has read-only permission!

#### Stacked Borrows rules

- Memory access: find permission for our tag, remove incompatible items above
- Assigning fresh tags:
  - Taking a reference (&mut term): fresh tag n. find old tag; remove incompatible above; push new
  - Reference to raw pointer (term as \*mut T): tag ⊥.
     find old tag; add new just above
- Compatibility:
  - Reads are compatible with SharedRW, SharedRO
  - Writes to SharedRW are compatible with SharedRW
- Retag when reference "enters" function (argument, load, call returns)

```
fn test_shared(x: &i32) -> bool {
  retag(x); // think: x = &*x;
  let val = *x;
  // unknown_function_2 cannot mutate x
  unknown_function_2(x);
  return *x == val; // must return true
}
```

x has fresh tag with permission SharedRO at the top of the stack

fn test\_shared(x: &i32) >> bool {
 retag(x); 
// think: x = &\*x;
let val = \*x;
// unknown function 2 connect muto

}

// unknown\_function\_2 cannot mutate x
unknown\_function\_2(x);

return \*x == val; // must return true

x has fresh tag with permission SharedRO at the top of the stack

fn test\_shared(x: &i32) bool {
 retag(x); // think: x = &\*x;

let val = \*x;

}

// unknown\_function\_2 cannot mutate x
unknown\_function\_2(x);

return \*x == val; // must return true

If unknown\_function\_2 writes to this memory, x's tag will be removed from the stack

x has fresh tag with permission SharedRO at the top of the stack

fn test\_shared(x: &i32) >> bool {
 retag(x);

let val = \*x;

// unknown\_function\_2 cannot mutate x
unknown\_function\_2(x);

return \*x == val; // must return true

UB unless x's permission is still in the stack

}

If unknown\_function\_2 writes to this memory, x's tag will be removed from the stack

#### What else?

#### What I didn't talk about:

- Interior mutability (shared references through which mutation is allowed)
- Barriers, two-phase borrows
- Formal model in Coq (proofs of optimizations in progress)

#### Future work:

- Integrating stacked borrows into RustBelt
- Handling integer-pointer casts
- Proving correctness of compilation to LLVM

For more details,
 check out Ralf's blog at:
https://www.ralfj.de/blog/